European Journal of Combinatorics 32 (2011) 1054–1083

Contents lists available at ScienceDirect

European Journal of Combinatorics

journal homepage: www.elsevier.com/locate/ejc

Quasi-random graphs and graph limits

Svante Janson

Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden

ARTICLE INFO

Article history:
Available online 17 April 2011

ABSTRACT

We use the theory of graph limits to study several quasi-random properties, mainly dealing with various versions of hereditary subgraph counts. The main idea is to transfer the properties of (sequences of) graphs to properties of graphons, and to show that the resulting graphon properties only can be satisfied by constant graphons. These quasi-random properties have been studied before by other authors, but our approach gives proofs that we find cleaner, and which avoid the error terms and ε in the traditional arguments using the Szemerédi regularity lemma. On the other hand, other technical problems sometimes arise in analysing the graphon properties; in particular, a measure-theoretic problem on elimination of null sets that arises in this way is treated in an appendix.

1. Introduction

A *quasi-random* graph is a graph that 'looks like' a random graph. Formally, this is best defined for a sequence of graphs (G_n) with $|G_n| \to \infty$. Thomason [22,23] and Chung et al. [7] showed that a number of different 'random-like' conditions on such a sequence are equivalent, and we say that (G_n) is p-quasi-random if it satisfies these conditions. (Here $p \in [0, 1]$ is a parameter.) We give one of these conditions, which is based on subgraph counts, in (2.1) below. Other characterisations have been added by various authors. The present paper studies in particular hereditarily extended subgraph count properties found by Simonovits and Sós [19,20], Shapira [16], Shapira and Yuster [17] and Yuster [24]; see Section 3. See also Sections 9 and 10 for further related equivalent properties (on sizes of cuts) found by Chung et al. [7] and Chung and Graham [6].

The theory of *graph limits* also concern the asymptotic behaviour of sequences (G_n) of graphs with $|G_n| \to \infty$. A notion of convergence of such sequences was introduced by Lovász and

E-mail address: svante.janson@math.uu.se.
URL: http://www.math.uu.se/~svante/.

0195-6698/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ejc.2011.03.011

Szegedy [14] and further developed by Borgs et al. [4,5]. This may be seen as giving the space of (unlabelled) graphs a suitable metric; the convergent sequences are the Cauchy sequences in this metric, and the completion of the space of unlabelled graphs in this metric is the space of (graphs and) graph limits. The graph limits are thus defined in a rather abstract way, but there are also more concrete representations of them. One important representation [14,4] uses a symmetric (Lebesgue) measurable function $W : [0, 1]^2 \to [0, 1]$; such a function is called a *graphon*, and defines a unique graph limit; see Section 2 for details. Note, however, that the representation is not unique; different graphons may be equivalent in the sense of defining the same graph limit. See further [3,8].

We write, with a minor abuse of notation, $G_n \to W$, if (G_n) is a sequence of graphs and W is a graphon such that (G_n) converges to the graph limit defined by W. It is well known that quasi-random graphs provide the simplest example of this: (G_n) is p-quasi-random if and only if $G_n \to p$, where p is the graphon that is constant p [14].

A central tool to study large dense graphs is Szemerédi's regularity lemma, and it is not surprising that this is closely connected to the theory of graph limits; see, e.g. , [4,15]. The Szemerédi regularity lemma is also important for the study of quasi-random graphs. For example, Simonovits and Sós [18] gave a characterisation of quasi-random graphs in terms of Szemerédi partitions. Moreover, the proofs in [19,20,16,17] that various properties characterise quasi-random graphs (see Section 3) use the Szemerédi regularity lemma. Roughly speaking, the idea is to take a Szemerédi partition of the graph and use the property to show that the Szemerédi partition has almost constant densities.

The main purpose of this paper is to point out that these, and other similar, characterisations of quasi-random graphs alternatively can be proved by replacing the Szemerédi regularity lemma and Szemerédi partitions by graph limit theory. The idea is to first take a graph limit of the sequence (or, in general, of a subsequence) and a representing graphon, then the property we assume for the graphs is translated into a property of the graphon, and finally it is proved that this graphon then has to be (a.e.) constant. We do this for several different related characterisations below. Our proofs will all have the same structure and consist of three parts, considering a sequence of graphs (G_n) and a graphon W with $G_n \to W$:

(i) An equivalence between a condition on subgraph counts in G_n and a corresponding condition for integrals of a functional Ψ of W. (Ψ is a function on $[0, 1]^m$ for some m, and is a polynomial in $W(x_i, x_j)$, $1 \le i < j \le m$.)

(ii) An equivalence between this integral condition on Ψ and a pointwise condition on Ψ.

(iii) An equivalence between this pointwise condition on Ψ and $W = p$.

In all the cases that we consider, (i) is rather straightforward, and performed in essentially the same way for all versions. Step (ii) follows from some version of the Lebesgue differentiation theorem, although some cases are more complicated than others. The arguments used in (iii) are similar to the arguments in earlier proofs that the Szemerédi partition has almost constant densities (under the corresponding condition on the graphs) and the algebraic problems that arise in some cases will be the same. However, the use of graph limits eliminates the many error terms and ε inherent in arguments using the Szemerédi regularity lemma, and provides at least sometimes proofs that are simpler and cleaner. With some simplification, we can say that we split the proofs into three parts (i)–(iii) which are combinatorial, analytic and algebraic, respectively. This has the advantage of isolating different types of technical difficulties; moreover, it allows us to reuse some steps that are the same for several different cases. (See for example Section 7 where we prove several variants of the characterisations by modifying step (i) or (ii).) On the other hand, it has to be admitted that there can be technical problems with the analysis of the graphons too, especially in (ii), and that our approach does not simplify the algebraic problems in (iii). (In particular, we have not been able to improve the results in [20], where it is this algebraic part that has not yet been done for general graphs.) The algebraic parts of the proofs are thus essentially the same as in previous proofs, but cleaner. Somewhat disappointingly, it seems that the graph limit method offers greatest simplifications in the simplest cases. At the end, it is partly a matter of taste if one prefers the finite arguments using the Szemerédi regularity lemma or the infinitesimal arguments using graphons; we invite the reader to make comparisons.

2. Preliminaries and notation

All graphs in this paper are finite, undirected and simple. The vertex and edge sets of a graph G are denoted by $V(G)$ and $E(G)$. We write $|G| := |V(G)|$ for the number of vertices of G, and $e(G) := |E(G)|$ for the number of edges. $\overline{G}$ is the complement of G. As usual, $[n] := \{1, \dots, n\}$.

2.1. Subgraph counts

Let F and G be graphs. It is convenient to assume that the graphs are labelled, with $V(F) = [|F|] := \{1, \dots, |F|\}$, but the labelling does not affect our results.

Definition 2.1. $N(F, G)$ is the number of labelled (not necessarily induced) copies of F in G; equivalently, $N(F, G)$ is the number of injective maps $\varphi : V(F) \to V(G)$ that are graph homomorphisms (i.e., if i and j are adjacent in F, then $\varphi(i)$ and $\varphi(j)$ are adjacent in G).

If U is a subset of $V(G)$, we further define $N(F, G; U)$ as the number of such copies with all vertices in U; thus $N(F, G; U) = N(F, G|_U)$.

More generally, if $U_1, \dots, U_{|F|}$ are subsets of $V(G)$, we define $N(F, G; U_1, \dots, U_{|F|})$ to be the number of labelled copies of F in G with the ith vertex in U_i; equivalently, $N(F, G; U_1, \dots, U_{|F|})$ is the number of injective graph homomorphisms $\varphi : F \to G$ such that $\varphi(i) \in U_i$ for every $i \in V(F)$. (Note that we consider a fixed labelling of the vertices of F and count the number of copies where vertex i is in U_i, so the labelling and the ordering of $U_1, \dots, U_{|F|}$ are important. However, our conditions below, such as (3.2), will be symmetric in $U_1, \dots, U_{|F|}$, so the order and the labelling are not important there.)

2.2. Quasi-random graphs

One of the several equivalent definitions of quasi-random graphs by Chung et al. [7] is:

Definition 2.2. A sequence of graphs (G_n) (with $|G_n| \to \infty$) is p-quasi-random if and only if, for every graph F,

$$N(F, G_n) = (p^{e(F)} + o(1))|G_n|^{|F|}. \tag{2.1}$$

(All unspecified limits in this paper are as $n \to \infty$, and $o(1)$ denotes a quantity that tends to 0 as $n \to \infty$. We will often use $o(1)$ for quantities that depend on some subset(s) of a vertex set $V(G)$ or of $[0, 1]$; we then always implicitly assume that the convergence is uniform for all choices of the subsets. We interpret $o(a_n)$ for a given sequence a_n similarly.)

It turns out that it is not necessary to require (2.1) for all graphs F; in particular, it suffices to use the graphs K_2 and C_4 [7]. However, it is not enough to require (2.1) for just one graph F. As a substitute, Sós [19] showed that a hereditary version of (2.1) for a single F is sufficient; see Section 3.

2.3. Graph limits

The graph limit theory is also based on the subgraph counts $N(F, G)$ (or the asymptotically equivalent number counting not necessarily injective graph homomorphisms $F \to G$, see [14,4]).

A sequence (G_n) of graphs, with $|G_n| \to \infty$, *converges*, if the numbers $t_{\mathrm{inj}}(F, G_n) := N(F, G_n)/(|G_n|)_{|F|}$ converge as $n \to \infty$, for every fixed graph F. (Here, $(|G_n|)_{|F|}$ denotes the falling factorial, which is the total number of injective maps $V(F) \to V(G_n)$, so $t_{\mathrm{inj}}(F, G_n)$ is the proportion of injective maps that are homomorphisms. Since we consider limits as $|G_n| \to \infty$ only, we could as well instead consider $t(F, G_n)$, the proportion of *all* maps $V(F) \to V(G_n)$ that are homomorphisms, or the hybrid version $N(F, G_n)/|G_n|^{|F|}$.) Note that the numbers $t_{\mathrm{inj}}(F, G_n) \in [0, 1]$, which implies the compactness property that every sequence (G_n) of graphs with $|G_n| \to \infty$ has a convergent subsequence. For details and several other equivalent properties, see [14,4,5,8].

The graph limits that arise in this way may be thought of as elements of a completion of the space of (unlabelled) graphs with a suitable metric. One useful representation [14,4] uses a symmetric measurable function $W : [0,1]^2 \to [0,1]$; such a function is called a *graphon*, and defines a graph limit in the following way. If F is a labelled graph and W a graphon, we define

$$\Psi_{F,W}(x_1,\dots,x_{|F|}) := \prod_{ij\in E(F)} W(x_i,x_j) \tag{2.2}$$

and (where the labelling does no longer matter)

$$t(F,W) := \int_{[0,1]^{|F|}} \Psi_{F,W}. \tag{2.3}$$

(All integrals in this paper are with respect to the Lebesgue measure in one or several dimensions, unless, in the Appendix, we specify another measure.) A sequence (G_n) converges to the graph limit defined by W if $|G_n| \to \infty$ and

$$\lim_{n\to\infty} t_{\mathrm{inj}}(F,G_n) = t(F,W) \tag{2.4}$$

(or, equivalently, $t(F,G_n) \to t(F,W)$) for every F; as said above, in this case we write $G_n \to W$, although it should be remembered that the representation of the limit by a graphon W is not unique.

Definition 2.3. Two graphons W_1 and W_2 are *equivalent* if they define the same graph limit. (Equivalently, $G_n \to W_1 \iff G_n \to W_2$. By (2.4), this is further equivalent to $t(F,W_1) = t(F,W_2)$ for every graph F.)

Further, W_2 is a *version* of W_1 if $W_1 = W_2$ a.e.

Trivially, any version of a graphon W is equivalent to W, but there are also other possibilities; for example, we may rearrange W as in (2.11) below. See [4,3,8,2] for details on the non-uniqueness and characterisations of equivalent graphons.

Condition (2.1) can be written $t_{\mathrm{inj}}(F,G_n) \to p^{e(F)}$. Since the constant graphon $W = p$ has $t(F,W) = p^{e(F)}$ for every F by (2.2)–(2.3), this shows that, as said in Section 1, (G_n) is p-quasi-random if and only if $G_n \to p$. Note that in this case there is no problem with non-uniqueness: W is equivalent to the constant graphon p if and only if it is a version of p, i.e., $W = p$ a.e. (This is, for example, a consequence of Lemma 4.2 below, taking $F = K_2$ and using (i) $\iff$ (iii) for both p and W.)

2.4. Graphons from graphs

If G is a graph, we define a corresponding graphon W_G by partitioning $[0,1]$ into $|G|$ intervals I_i of equal lengths $1/|G|$; we then define W_G to be 1 on every $I_i \times I_j$ such that $ij \in E(G)$, and 0 otherwise. It is easily seen that if G is a graph, then

$$N(F,G) = |G|^{|F|} \int_{[0,1]^{|F|}} \Psi_{F,W_G} + O\big(|G|^{|F|-1}\big). \tag{2.5}$$

(The error term is because we have chosen to count injective homomorphisms only, cf. [14,4].) More generally, if $U_1,\dots,U_{|F|}$ are subsets of $V(G)$ and $U_1',\dots,U_{|F|}'$ are the corresponding subsets of $[0,1]$ given by $U_i' := \bigcup_{j\in U_i} I_j$, then

$$N(F,G;U_1,\dots,U_{|F|}) = |G|^{|F|} \int_{U_1'\times\cdots\times U_{|F|}'} \Psi_{F,W_G} + O\big(|G|^{|F|-1}\big). \tag{2.6}$$

2.5. Induced subgraph counts

Consider again labelled graphs F and G. We make the following definitions in analogy with Section 2.1.

Definition 2.4. $N^*(F, G)$ is the number of *induced* labelled copies of F in G; equivalently, $N^*(F, G)$ is the number of injective maps $\varphi : V(F) \to V(G)$ such that i and j are adjacent in $F \iff \varphi(i)$ and $\varphi(j)$ are adjacent in G.

Further, $N^*(F, G; U)$ is the number of such copies with all vertices in U and $N^*(F, G; U_1, \dots, U_{|F|})$ is the number of induced labelled copies of F in G with the ith vertex in U_i. (Here $U, U_1, \dots, U_{|F|} \subseteq V(G)$. Again, we consider a fixed given labelling of the vertices of F so the ordering of $U_1, \dots, U_{|F|}$ are in general important, but in our applications as (3.5), the labelling and ordering do not matter.)

For a graphon W we make the corresponding definitions, cf. Section 2.3,

$$\Psi^*_{F,W}(x_1, \dots, x_{|F|}) := \prod_{ij\in E(F)} W(x_i, x_j) \prod_{ij\notin E(F)} \bigl(1 - W(x_i, x_j)\bigr) \tag{2.7}$$

and

$$t_{\mathrm{ind}}(F, W) := \int_{[0,1]^{|F|}} \Psi^*_{F,W}. \tag{2.8}$$

Then, for any graph G, in analogy with (2.6) and using the notation there,

$$N^*(F, G; U_1, \dots, U_{|F|}) = |G|^{|F|} \int_{U'_1\times\cdots\times U'_{|F|}} \Psi^*_{F,W_G} + O\bigl(|G|^{|F|-1}\bigr). \tag{2.9}$$

Remark 2.5. If we define $t_{\mathrm{ind}}(F, G) := N^*(F, G)/(|G|)_{|F|}$, then the convergence criterion (2.4) (for every F) is equivalent to $t_{\mathrm{ind}}(F, G_n) \to t_{\mathrm{ind}}(F, W)$ (for every F) by inclusion–exclusion [14,4].

2.6. Cut norm and cut metric

Definition 2.6. The *cut norm* $\|W\|_\square$ of $W \in L^1([0, 1]^2)$ is defined by

$$\|W\|_\square := \sup_{S,T\subseteq[0,1]} \left| \int_{S\times T} W(x, y)\,\mathrm{d}x\,\mathrm{d}y \right|. \tag{2.10}$$

A *rearrangement* of the graphon W is any graphon W^φ defined by

$$W^\varphi(x, y) = W(\varphi(x), \varphi(y)), \tag{2.11}$$

where $\varphi : [0, 1] \to [0, 1]$ is a measure preserving bijection.

The *cut metric* δ by Borgs et al. [4] may be defined by, for two graphons W_1, W_2,

$$\delta_\square(W_1, W_2) = \inf_\varphi \|W_1 - W_2^\varphi\|_\square, \tag{2.12}$$

where the infimum is over all rearrangements of W_2. (It makes no difference if we rearrange W_1 instead, or both W_1 and W_2.)

A major result of Borgs et al. [4] is that if $|G_n| \to \infty$, then $G_n \to W \iff \delta_\square(W_{G_n}, W) \to 0$, so convergence of a sequence of graphs as defined above is the same as convergence in the metric $\delta_\square$.

3. Subgraph counts in induced subgraphs

Simonovits and Sós [19] gave the following characterisation of p-quasi-random graphs using the number of subgraphs of a given type in induced subgraphs. (The case $F = K_2$, when $N(K_2, G_n; U)$ is twice the number of edges with both endpoints in U, is one of the original quasi-random properties in [7].)

Theorem 3.1 (*Simonovits and Sós [19]*). *Suppose that (G_n) is a sequence of graphs with $|G_n| \to \infty$. Let F be any fixed graph with $e(F) > 0$ and let $0 < p \le 1$. Then (G_n) is p-quasi-random if and only if, for all subsets U of $V(G_n)$,*

$$N(F, G_n; U) = p^{e(F)}|U|^{|F|} + o\big(|G_n|^{|F|}\big). \tag{3.1}$$

For our discussion of graph limit method, it is also interesting to consider the following weaker version (with a stronger hypothesis), patterned after Theorem 3.11 below.

Theorem 3.2. *Suppose that (G_n) is a sequence of graphs with $|G_n| \to \infty$. Let F be any fixed graph with $e(F) > 0$ and let $0 < p \le 1$. Then (G_n) is p-quasi-random if and only if, for all subsets $U_1, \dots, U_{|F|}$ of $V(G_n)$,*

$$N(F, G_n; U_1, \dots, U_{|F|}) = p^{e(F)} \prod_{i=1}^{|F|} |U_i| + o\big(|G_n|^{|F|}\big). \tag{3.2}$$

Graph limit proofs of these and other theorems in this section are given in Sections 4–8.

Remark 3.3. Since (3.1) is the special case of (3.2) with $U_1 = \cdots = U_{|F|}$, the 'if' direction of Theorem 3.2 is a corollary of Theorem 3.1. The 'only if' direction does not follow immediately from Theorem 3.1, but it is straightforward to prove, either by the methods of [19] or by our methods with graph limits, see Section 4; hence the main interest is in the 'if' direction. (The same is true for the results below for the induced case.)

Remark 3.4. Theorems 3.1 and 3.2 obviously fail when $e(F) = 0$, since then (3.1) and (3.2) hold trivially and the assumptions give no information on G_n. They fail also if $p = 0$; for example, if $F = K_3$ and G_n is the complete bipartite graph $K_{n,n}$.

Shapira [16] and Shapira and Yuster [17] consider also an intermediate version where a symmetric form of (3.2) is used, summing over all permutations of $(U_1, \dots, U_{|F|})$ (or, equivalently, over all labellings of F); moreover, $U_1, \dots, U_{|F|}$ are supposed to be disjoint and of the same size. It is shown directly in [16] that this is equivalent to (3.1). See also Sections 7.1 and 7.2.

The main result of Shapira [16] is that Theorem 3.1 remains valid even if we only require (3.1) for U of size $\alpha|G_n|$ with $\alpha = 1/(|F| + 1)$. (It is a simple consequence that any smaller positive α will also do.) This was improved by Yuster [24], who proved this for any $\alpha \in (0, 1)$. We state this, and the corresponding result for a sequence of (disjoint) subsets.

Theorem 3.5 (*Yuster [24]*). *Let (G_n), F and p be as in Theorem 3.1, and let $0 < \alpha < 1$. Then (G_n) is p-quasi-random if and only if (3.1) holds for all subsets U of $V(G_n)$ with $|U| = \lfloor \alpha|G_n| \rfloor$.*

Theorem 3.6. *Let (G_n), F and p be as in Theorem 3.2, and let $0 < \alpha < 1$. Then (G_n) is p-quasi-random if and only if (3.2) holds for all subsets $U_1, \dots, U_{|F|}$ of $V(G_n)$ with $|U_i| = \lfloor \alpha|G_n| \rfloor$.*

If $\alpha < 1/|F|$, it is enough to assume (3.2) for $U_1, \dots, U_{|F|}$ that are further disjoint.

For $F = K_2$, Theorem 3.5 with $\alpha = 1/2$ is another of the original characterisations by Chung et al. [7], and the generalisation to arbitrary $\alpha \in (0, 1)$ is stated in [6]. Another related characterisation from [6] is discussed in Section 9.

Turning to induced copies of F, the situation is much more complicated, as discussed in Simonovits and Sós [20]. First, the expected number of induced labelled copies of F in a random graph $G(n, p)$ is $\beta_F(p)n^{|F|} + o(n^{|F|})$, with

$$\beta_F(p) := p^{e(F)}(1-p)^{e(\overline{F})} = p^{e(F)}(1-p)^{\binom{|F|}{2} - e(F)}. \tag{3.3}$$

Hence, the condition corresponding to (3.1) for induced subgraphs is: For all subsets U of $V(G_n)$,

$$N^*(F, G_n; U) = \beta_F(p)|U|^{|F|} + o\big(|G_n|^{|F|}\big). \tag{3.4}$$

Indeed, as observed in [19,20], this holds for every p-quasi-random (G_n), but the converse is generally false. One reason is that, provided F is neither empty nor complete, then $\beta_F(0) = \beta_F(1) = 0$, and if $p_F := e(F)/\binom{|F|}{2}$ (the edge density in F), then $\beta_F(p)$ increases on $[0, p_F]$ and decreases on $[p_F, 1]$. Hence, for every $p \neq p_F$, there is another $\bar{p}$ such that $\beta_F(\bar{p}) = \beta_F(p)$; we call p and $\bar{p}$ *conjugate*. (For completeness, we let $\bar{p} := p$ when $p = p_F$ or when F is empty or complete. Note also that $\bar{p}$ depends on F as well as p.) Obviously, a $\bar{p}$-quasi-random sequence (G_n) also satisfies (3.4). Moreover, any combination of a p-quasi-random sequence and a $\bar{p}$-quasi-random sequence will satisfy (3.4). Hence the best we can hope for is the following. We say that (G_n) is *mixed* $(p, \bar{p})$*-quasi-random* if it is p-quasi-random, $\bar{p}$-quasi-random, or a combination of two such sequences.

Definition 3.7. Let $0 \le p \le 1$. We say that a graph F is *hereditary induced forcing* ($\mathsf{HI}(p)$) if every (G_n) that satisfies (3.4) for all subsets U of $V(G_n)$ is mixed $(p, \bar{p})$-quasi-random. In this case we also write $F \in \mathsf{HI}(p)$ (thus regarding $\mathsf{HI}(p)$ as a set of graphs).

We say that F is HI (and write $F \in \mathsf{HI}$) if F is $\mathsf{HI}(p)$ for every $p \in (0, 1)$ (thus excluding the rather exceptional cases $p = 0$ and $p = 1$).

Remark 3.8. The definition of mixed $(p, \bar{p})$-quasi-random is perhaps better stated in terms of graph limits. Just as (G_n) is p-quasi-random if and only if $G_n \to p$, where p stands for the graphon that is constant p, (G_n) is mixed $(p, \bar{p})$-quasi-random if and only if the limit points of (G_n) are contained in $\{p, \bar{p}\}$, i.e., if every convergent subsequence of (G_n) converges to either the graphon p or the graphon $\bar{p}$.

In general we say that a sequence (G_n), with $|G_n| \to \infty$ as always, is *mixed quasi-random* if the set of limit points is contained in $\{p : p \in [0, 1]\}$, i.e., if every convergent subsequence converges to a constant graphon. (Equivalently, if every convergent subsequence is quasi-random).

Remark 3.9. Just as one talks about quasi-random properties of graphs, or more properly of sequences (G_n) of graphs, we say that a property of graphons W is *p-quasi-random* if it is satisfied only by $W = p$ a.e., that it is *quasi-random* if it is p-quasi-random for some $p \in [0, 1]$, and that it is *mixed quasi-random* if it is satisfied only by graphons that are a.e. constant (for some set of accepted constants).

Simonovits and Sós [20] gave a counter-example showing that the path P_3 with 3 vertices is *not* HI. They also showed that every regular F (with $|F| \ge 2$) is HI, and conjectured that P_3 and its complement $\overline{P}_3$ are the only graphs not in HI. This conjecture remains open. (The methods of the present paper do not seem to help.)

Remark 3.10. The cases F empty or complete are exceptional and rather trivial. If F is a complete graph K_m ($m \ge 2$), then $N^*(F, G_n; U) = N(F, G_n; U)$, and thus (3.4) implies that (G_n) is p-quasi-random by Theorem 3.1 (but not for $p = 0$ unless $m = 2$, see Remark 3.4). By taking complements we see that the same holds for an empty graph E_m ($m \ge 2$) and $0 \le p < 1$.

In particular, $E_m, K_m \in \mathsf{HI}$ when $m \ge 2$.

In view of the fact that not all graphs are HI, Shapira and Yuster [17] gave the following substitute, which is an induced version of Theorem 3.2.

Theorem 3.11 (*Shapira and Yuster [17]*)**.** *Suppose that (G_n) is a sequence of graphs with $|G_n| \to \infty$. Let F be any fixed graph with $|F| > 1$ and let $0 < p < 1$. Then (G_n) is mixed $(p, \bar{p})$-quasi-random if and only if, for all subsets $U_1, \dots, U_{|F|}$ of $V(G_n)$,*

$$N^*(F, G_n; U_1, \dots, U_{|F|}) = p^{e(F)}(1-p)^{\binom{|F|}{2} - e(F)} \prod_{i=1}^{|F|} |U_i| + o\big(|G_n|^{|F|}\big). \tag{3.5}$$

Moreover, it suffices that (3.5) holds for all sequences $U_1, \dots, U_{|F|}$ of disjoint subsets of $V(G_n)$ with the same size, $|U_1| = \dots = |U_{|F|}|$.

To show the flexibility with which our method combines different conditions, we also show that it suffices to consider subsets of a given size for induced subgraph counts too, in analogy with Theorems 3.5 and 3.6.

Theorem 3.12. *In Theorem 3.11, it suffices that (3.5) holds for all sequences $U_1, \dots, U_{|F|}$ of subsets of $V(G_n)$ with $|U_i| = \lfloor \alpha |G_n| \rfloor$, for any fixed α with $0 < \alpha < 1$. Alternatively, if $0 < \alpha < 1/|F|$, it suffices that (3.5) holds for all such sequences of disjoint $U_1, \dots, U_{|F|}$.*

Theorem 3.13. *Let $0 < \alpha < 1$ and $0 \le p \le 1$, and let F be a fixed graph with $F \in \mathsf{HI}(p)$. Then every sequence (G_n) with $|G_n| \to \infty$ such that (3.4) holds for all subsets U of $V(G_n)$ with $|U| = \lfloor \alpha |G_n| \rfloor$ is mixed $(p, \bar{p})$-quasi-random.*

Remark 3.14. Theorems 3.6 and 3.12 fail for disjoint sets $U_1, \dots, U_{|F|}$ in the limiting case $\alpha = 1/|F|$, at least for $F = K_2$; see Section 9 and Remark 7.4. We leave it as an open problem to investigate this case for other graphs F.

4. Graph limit proof of Theorem 3.2

We give proofs of the theorems above using graph limits; the reader should compare these to the combinatorial proofs in [19,20,16,17,24] using the Szemerédi regularity lemma. In order to exhibit the main ideas clearly, we begin in this section with the simplest case and give a detailed proof of Theorem 3.2. In fact, as a preparation for later results we give two versions of the proof that differ in the treatment of a technical problem, and in the next section we give further variations of the proof. In the following sections we will give the minor modifications needed for the other results, treating the additional complications one by one.

The first step is to recall that the space of graphs and graph limits is compact; thus, every sequence has a convergent subsequence [4]. Hence, if (G_n) is not p-quasi-random, we can select a subsequence (which we also denote by (G_n)), such that $G_n \to W$ for some graphon W that is not equivalent to the constant graphon p, which simply means that $W \neq p$ on a set of positive measure.

Hence, in order to prove Theorem 3.2, it suffices to assume further that $G_n \to W$ for some graphon W, and then prove that $W = p$ a.e.

4.1. Translating to graphons

In this subsection we use the graph limit theory in [4] to translate property (3.2) to graph limits.

We begin with an easy consequence of the Lebesgue differentiation theorem; for future reference we state that it as a (well-known) lemma. (See Lemma 7.3 below for a stronger version.) We let λ denote the Lebesgue measure (in one or several dimensions).

Lemma 4.1. *Suppose that $f : [0, 1]^m \to \mathbb{R}$ is an integrable function such that $\int_{A_1 \times \cdots \times A_m} f = 0$ for all sequences $A_1, \dots, A_m$ of disjoint measurable subsets of $[0, 1]$. Then $f = 0$ a.e.*

Moreover, it is enough to consider $A_1, \dots, A_m$ with $\lambda(A_1) = \cdots = \lambda(A_m)$; we may even further impose that $\lambda(A_k) \in \{\varepsilon_1, \varepsilon_2, \dots\}$ for any given sequence $\varepsilon_n \to 0$.

Proof. For any distinct $x_1, \dots, x_m \in (0, 1)$ and any sufficiently small $\varepsilon > 0$ we take $A_i = (x_i - \varepsilon, x_i + \varepsilon)$ and find

$$(2\varepsilon)^{-m} \int_{|y_i - x_i| < \varepsilon, i=1,\dots,m} f(y_1, \dots, y_m) = (2\varepsilon)^{-m} \int_{A_1 \times \cdots \times A_m} f = 0.$$

By the Lebesgue differentiation theorem, see e.g. [21, Section 1.8], the left-hand side converges to $f(x_1, \dots, x_m)$ as $\varepsilon \to 0$ for a.e. $x_1, \dots, x_m$. □

We can now easily translate condition (3.2) in Theorem 3.2 to a corresponding condition for the limiting graphon (which we may assume exists, as discussed above).

Lemma 4.2. *Suppose that $G_n \to W$ for some graphon W and let F be a fixed graph and $\gamma \geq 0$ a fixed number. Then the following are equivalent:*

(i) *For all subsets $U_1, \dots, U_{|F|}$ of $V(G_n)$,*

$$N(F, G_n; U_1, \dots, U_{|F|}) = \gamma \prod_{i=1}^{|F|} |U_i| + o\big(|G_n|^{|F|}\big). \tag{4.1}$$

(ii) *For all subsets $A_1, \dots, A_{|F|}$ of $[0, 1]$,*

$$\int_{A_1 \times \cdots \times A_{|F|}} \Psi_{F,W}(x_1, \dots, x_{|F|}) = \gamma \prod_{i=1}^{|F|} \lambda(A_i). \tag{4.2}$$

(iii) *$\Psi_{F,W}(x_1, \dots, x_{|F|}) = \gamma$ for a.e. $x_1, \dots, x_{|F|} \in [0, 1]^{|F|}$.*

Proof. (iii) $\Rightarrow$ (ii) is trivial, and (ii) $\Rightarrow$ (iii) is immediate by Lemma 4.1 applied to $\Psi_{F,W} - \gamma$.
(i) $\iff$ (ii). The convergence $G_n \to W$ is equivalent to $\delta_\square(W_{G_n}, W) \to 0$. By the definition of $\delta_\square$, there thus exist measure preserving bijections $\varphi_n : [0, 1] \to [0, 1]$ such that if $W_n := W_{G_n}^{\varphi_n}$, then $\|W_n - W\|_\square \to 0$. Fix n, and let I_{nj} $(1 \leq j \leq n)$ be the intervals of length $|G_n|^{-1}$ used to define W_{G_n}, and let as in (2.6) $U' := \bigcup_{j \in U} I_{nj}$ for a subset U of $V(G_n)$; further, let $I''_{nj} := \varphi_n^{-1}(I_{nj})$ and $U'' := \varphi_n^{-1}(U') = \bigcup_{j \in U} I''_{nj}$. Then, for any subsets $U_1, \dots, U_{|F|}$ of $V(G_n)$, by (2.6) and a change of variables,

$$N(F, G_n; U_1, \dots, U_{|F|}) = |G_n|^{|F|} \int_{U''_1 \times \cdots \times U''_{|F|}} \Psi_{F,W_n} + o\big(|G_n|^{|F|}\big).$$

Hence, (i) is equivalent to

$$\int_{U''_1 \times \cdots \times U''_{|F|}} \Psi_{F,W_n} = \gamma \prod_{i=1}^{|F|} \frac{|U_i|}{|G_n|} + o(1) = \gamma \prod_{i=1}^{|F|} \lambda(U''_i) + o(1), \tag{4.3}$$

for all subsets U''_i that are unions of sets I''_{nj}.

We next extend (4.3) from the special sets U''_i (in a family that depends on n) to arbitrary (measurable) sets. Thus, assume that (4.3) holds, and let $A_1, \dots, A_{|F|}$ be arbitrary subsets of $[0, 1]$. Fix n and let $a_{ij} := \lambda(A_i \cap I''_{nj})/\lambda(I''_{nj})$. Further, let B_i be a random subset of $[0, 1]$ obtained by taking an independent family J_{ij} of independent 0–1 random variables with $\mathbb{P}(J_{ij} = 1) = a_{ij}$, and then taking $B_i := \bigcup_{j:J_{ij}=1} I''_{nj}$. Then the sets B_i are of the form U''_i, so (4.3) applies to them, and, noting that W_n is constant on every set $I''_{ni} \times I''_{nj}$, and hence Ψ_{F,W_n} is constant on every set $I''_{nj_1} \times \cdots \times I''_{nj_{|F|}}$,

$$\begin{aligned}
\int_{A_1 \times \cdots \times A_{|F|}} (\Psi_{F,W_n} - \gamma) &= \sum_{j_1, \dots, j_{|F|}=1}^{|G_n|} \prod_{i=1}^{|F|} a_{ij_i} \int_{I''_{nj_1} \times \cdots \times I''_{nj_{|F|}}} (\Psi_{F,W_n} - \gamma) \\
&= \mathbb{E} \sum_{j_1, \dots, j_{|F|}=1}^{|G_n|} \prod_{i=1}^{|F|} J_{ij_i} \int_{I''_{nj_1} \times \cdots \times I''_{nj_{|F|}}} (\Psi_{F,W_n} - \gamma) \\
&= \mathbb{E} \int_{B_1 \times \cdots \times B_{|F|}} (\Psi_{F,W_n} - \gamma) = o(1),
\end{aligned} \tag{4.4}$$

where the final estimate uses (4.3). Consequently, (4.3), for all special sets U''_i, is equivalent to the same estimate

$$\int_{A_1 \times \cdots \times A_{|F|}} \Psi_{F,W_n} = \gamma \prod_{i=1}^{|F|} \lambda(A_i) + o(1), \tag{4.5}$$

for any measurable sets $A_1, \dots, A_{|F|}$ in $[0, 1]$. Consequently, (i) is equivalent to (4.5). (Recall that estimates such as (4.5) are supposed to be uniform over all choices of $A_1, \dots, A_{|F|}$.)

It is well known that for two graphons W and W',

$$\left| \int_{[0,1]^m} \big(\Psi_{F,W} - \Psi_{F,W'}\big) \right| = O(\|W - W'\|_\square),$$

see [4]; moreover, the proof in [4] (or the version of the proof in [2]) shows that the same holds, uniformly, also if we integrate over a subset $A_1 \times \cdots \times A_m$. (In other words, extending the cut norm to functions of several variables as in [1], $\|\Psi_{F,W} - \Psi_{F,W'}\|_\square = O(\|W - W'\|_\square)$.) Consequently, the assumption $G_n \to W$, which as said yields $\|W_n - W\|_\square \to 0$, implies that $\int_{A_1\times\cdots\times A_{|F|}} \Psi_{F,W_n} = \int_{A_1\times\cdots\times A_{|F|}} \Psi_{F,W} + o(1)$, and thus (4.5), and hence (i), is equivalent to

$$\int_{A_1\times\cdots\times A_{|F|}} \Psi_{F,W} = \gamma \prod_{i=1}^{|F|} \lambda(A_i) + o(1). \tag{4.6}$$

Consequently, (ii) $\Rightarrow$ (i). Conversely, none of the terms in (4.6) depends on n, so if (4.6) holds, then the $o(1)$ error term vanishes and (4.2) holds. Hence (i) $\Rightarrow$ (ii). □

4.2. *An optional measure-theoretic interlude*

To prove Theorem 3.2, it thus remains only to show that if W is a graphon such that $\Psi_{F,W} = p^{e(F)}$ a.e., then $W = p$ a.e. (In the terminology of Remark 3.9, "$\Psi_{F,W} = p^{e(F)}$" is a p-quasi-random property.)

We know several ways to do this. One, direct, is given in Section 4.4. However, as will be seen in Section 4.3, it is much simpler to argue if we can assume that $\Psi_{F,W} = p^{e(F)}$ everywhere, and not just a.e. (The main reason is that we then can choose $x_1 = x_2 = \cdots = x_{|F|}$.) Hence, somewhat surprisingly, the qualification 'a.e.' here forms a significant technical problem. Usually, 'a.e.' is just a technical formality in arguments in integration and measure theory, but here it is an obstacle and we would like to get rid of it. We do not see any trivial way to do this, but we can do it as follows. (Recall from Definition 2.3 that a version of W is a graphon that is a.e. equal to W; this implies that W' is equivalent to W and thus $G_n \to W'$ as well.) See Section 5 for a more general, but weaker, result.

Lemma 4.3. *Let F be a graph with $e(F) > 0$, and let W be a graphon. If $\Psi_{F,W} = \gamma > 0$ a.e. on $[0, 1]^{|F|}$, then there exists a version W' of W such that $\Psi_{F,W'}(x_1, \dots, x_{|F|}) = \gamma$ for all $(x_1, \dots, x_{|F|}) \in [0, 1]^{|F|}$.*

Proof. By symmetry, we may assume that $12 \in E(F)$; hence $\Psi_{F,W}(x_1, \dots, x_{|F|})$, defined in (2.2), contains a factor $W(x_1, x_2)$. We let $x' := (x_3, \dots, x_{|F|})$ and collect the other factors in (2.2) into a product $f(x_1, x')$ of the factors corresponding to edges $1j \in E(F)$ with $j \ge 3$, and another product $g(x_2, x')$ of the remaining factors. Thus

$$\Psi_{F,W}(x_1, \dots, x_{|F|}) = W(x_1, x_2) f(x_1, x') g(x_2, x').$$

By assumption, thus

$$W(x_1, x_2) f(x_1, x') g(x_2, x') = \gamma \tag{4.7}$$

for a.e. (x_1, x_2, x'). We may thus choose x' (a.e. choice will do) such that (4.7) holds for a.e. (x_1, x_2). We fix one such x' and write $f(x) := f(x, x')$, $g(y) := g(y, x')$; we then have $W(x, y) f(x) g(y) = \gamma$ for a.e. (x, y).

We define $W_1(x, y) := \max\big(1, \gamma/(f(x)g(y))\big)$; thus $W_1 = W$ a.e.

Let $|(x_1, \dots, x_m)|_\infty := \max |x_i|$. Recall that if f is an integrable function on $\mathbb{R}^m$ for some m (or on a subset such as $[0, 1]^m$), then a point x is a *Lebesgue point* of f if $(2\varepsilon)^{-m} \int_{|y-x|_\infty<\varepsilon} |f(y) - f(x)|\,dy = o(1)$ as $\varepsilon \to 0$. In probabilistic terms, this says that if X_x^ε is a random point in the cube $\{y : |y - x|_\infty < \varepsilon\}$, then $f(X_x^\varepsilon) \xrightarrow{L^1} f(x)$. For bounded functions, which is the case here, this is equivalent to $f(X_x^\varepsilon) \xrightarrow{\mathrm{p}} f(x)$ as $\varepsilon \to 0$, which shows, for example, that if x is a Lebesgue point of both f and g, then it is also a Lebesgue point of $f \pm g$, fg, and, provided $g(x) \neq 0$, of f/g. It is well known, see e.g. Stein [21, Section 1.8], that if f is integrable, then a.e. point is a Lebesgue point of f.

We can thus find a null set $N \subset [0, 1]$ such that every $x \in \mathcal{S} := [0, 1] \setminus N$ is a Lebesgue point of both f and g. Since $W(x, y) \leq 1$ and thus $f(x)g(y) \geq \gamma$ a.e., it then follows that if $(x_1, x_2) \in \mathcal{S}^2$, then (x_1, x_2) is a Lebesgue point of W_1. This implies, by the definition (2.2), that if $(x_1, \dots, x_{|F|}) \in \mathcal{S}^{|F|}$, then $(x_1, \dots, x_{|F|})$ is a Lebesgue point of $\Psi_{F,W}$; hence, using $\Psi_{F,W_1} = \Psi_{F,W}$ a.e. and $\Psi_{F,W} = \gamma$ a.e., $\Psi_{F,W_1}(x_1, \dots, x_{|F|}) = \gamma$ for $(x_1, \dots, x_{|F|}) \in \mathcal{S}^{|F|}$.

This would really be enough for our purposes, but to obtain the conclusion as stated, we choose $x_0 \in \mathcal{S}$ and define $\varphi : [0, 1] \to [0, 1]$ by $\varphi(x) = x$ for $x \in \mathcal{S}$ and $\varphi(x) = x_0$ for $x \in N$; then $W' := W_1^{\varphi}$ satisfies $\Psi_{F,W'} = \gamma$ everywhere. □

Remark 4.4. Although we do not need it, we note that Lemma 4.3 is valid for the trivial case $e(F) = 0$ too, since then $\Psi_{F,W} = 1$ for every W and there is nothing to prove. We do not know whether Lemma 4.3 is also valid for $\gamma = 0$; consider for example $F = K_3$. (In this case it suffices to consider 0/1-valued W and W'.)

4.3. The first algebraic argument

The proof of Theorem 3.2 is now completed, by Lemmas 4.2 and 4.3 and the remarks above, by the following lemma:

Lemma 4.5. *Let F be a graph with $e(F) > 0$ and let W be a graphon. If $p > 0$ and $\Psi_{F,W}(x_1, \dots, x_{|F|}) = p^{e(F)}$ for every $(x_1, \dots, x_{|F|}) \in [0, 1]^{|F|}$, then $W = p$.*

Proof. First take $x_1 = x_2 = \cdots = x_{|F|} = x$. Then $\Psi_{F,W}(x_1, \dots, x_{|F|}) = W(x, x)^{e(F)}$, and thus $W(x, x) = p$, for every $x \in [0, 1]$. Next, we may assume by symmetry that the degree d_1 of vertex 1 in F is non-zero. Let $x, y \in [0, 1]$ and take $x_1 = x$ and $x_2 = \cdots = x_{|F|} = y$. Then

$$p^{e(F)} = \Psi_{F,W}(x_1, \dots, x_{|F|}) = W(x, y)^{d_1} W(y, y)^{e(F)-d_1} = W(x, y)^{d_1} p^{e(F)-d_1}.$$

Hence $W(x, y) = p$. □

This completes the first version of our graph limit proof of Theorem 3.2.

4.4. The second algebraic argument

As said above, we can alternatively avoid Lemma 4.3 and instead use the following stronger version of Lemma 4.5, which together with Lemma 4.2 yields another proof of Theorem 3.2.

Lemma 4.6. *Let F be a graph with $e(F) > 0$ and let W be a graphon. If $\Psi_{F,W}(x_1, \dots, x_{|F|}) = p^{e(F)}$ for a.e. $(x_1, \dots, x_{|F|}) \in [0, 1]^{|F|}$, then $W = p$ a.e.*

Proof. We first symmetrise. (Recall that $\Psi_{F,W}$ is defined by (2.2) for a fixed labelling of $x_1, \dots, x_{|F|}$, and is in general not symmetric in the variables.) If $\sigma \in \mathfrak{S}_{|F|}$, the symmetric group of all permutations of $\{1, \dots, |F|\}$, let $\sigma(F)$ be the image of F, with edges $\sigma(i)\sigma(j)$ for $ij \in E(F)$, and consider

$$\prod_{\sigma \in \mathfrak{S}_{|F|}} \Psi_{\sigma(F),W}(x_1, \dots, x_{|F|}) = \prod_{1 \leq i < j \leq |F|} W(x_i, x_j)^{e(F)k!/\binom{k}{2}},$$

where the equality follows because, by symmetry, each ij is an edge in $\sigma(F)$ for $e(F)k!/\binom{k}{2}$ permutations σ. By the assumption, this equals $p^{e(F)k!}$ a.e., so taking logarithms and dividing by $e(F)k!$ we obtain

$$\binom{k}{2}^{-1} \sum_{1 \leq i < j \leq |F|} \log W(x_i, x_j) = \log p, \quad \text{a.e.}$$

For a.e. $(x_1, \dots, x_{|F|+2})$, this holds for every subsequence of $|F|$ elements x_i; it then follows by Lemma 4.7 below, with $d = 2$, $h = |F|$ and $a(\{i,j\}) = \log W(x_i, x_j) - \log p$, that in this case $W(x_1, x_2) = p$. Hence $W(x_1, x_2) = p$ for a.e. (x_1, x_2). □

Lemma 4.7. *Suppose that $1 \le d \le h$, and let $a(I)$ be an array defined for all d-subsets I of $[h+d]$. Suppose further that for every h-subset J of $[h+d]$,*

$$\sum_{I \subseteq J} a(I) = 0, \tag{4.8}$$

summing over the $\binom{h}{d}$ subsets of size d. Then $a(I) = 0$ for every I.

Proof. This is a form of a result by Gottlieb [11]. (It is easily proved by fixing a d-subset I_0 and then summing (4.8) for all J with $|J \cap I_0| = k$, for $k = 0, \dots, d$; we omit the details.) □

5. A general measure-theoretic lemma

In Section 4.4 we gave a proof of Theorem 3.2 avoiding Lemma 4.3. For other results proved in the following sections, we do not know any similar proofs, and we use a method similar to the first proof above. This requires a more general version of Lemma 4.3.

A *multiaffine* polynomial is a polynomial in several variables $\{x_\nu\}_{\nu \in \mathcal{I}}$, for some (finite) index set $\mathcal{I}$, such that each variable has degree at most 1; it can thus be written as a linear combination of the $2^{|\mathcal{I}|}$ monomials $\prod_{\nu \in \mathcal{J}} x_\nu$ for subsets $\mathcal{J} \subseteq \mathcal{I}$. We are interested in the case when the index set $\mathcal{I}$ consists of the $\binom{m}{2}$ pairs $\{i,j\}$ with $1 \le i < j \le m$, for some $m \ge 2$. In this case we define, for any symmetric function $W : [0,1]^2 \to \mathbb{R}$ and $x_1, \dots, x_m \in [0,1]$,

$$\Phi_W(x_1, \dots, x_m) := \Phi\big((W(x_i, x_j))_{i<j}\big). \tag{5.1}$$

The functions $\Psi_{F,W}$ and $\Psi^*_{F,W}$ considered above are of this type; see (2.2) and (2.7), as well as their symmetrisations $\widetilde{\Psi}_{F,W}$ and $\widetilde{\Psi}^*_{F,W}$ considered below. In all our proofs we derive as an intermediate result an equation of the type $\Phi_W(x_1, \dots, x_m) = \gamma$ a.e. for some multiaffine Φ, and it would simplify the analysis of this equation if we were able to strengthen this to $\Phi_W(x_1, \dots, x_m) = \gamma$ for *every* $x_1, \dots, x_m \in [0,1]$, possibly after modifying W on a null set. We thus are led to the following measure-theoretic problem, with applications to quasi-random graphs:

Problem 5.1. Suppose that $\Phi\big((w_{ij})_{i<j}\big)$ is a multiaffine polynomial in the $\binom{m}{2}$ variables w_{ij}, $1 \le i < j \le m$, for some $m \ge 2$. Suppose further that $W : [0,1]^2 \to [0,1]$ is a graphon such that $\Phi_W(x_1, \dots, x_m) = \gamma$ a.e. for some $\gamma \in \mathbb{R}$. Does there always exist a graphon W' with $W' = W$ a.e. such that $\Phi_{W'}(x_1, \dots, x_m) = \gamma$ for *every* $x_1, \dots, x_m \in [0,1]$?

We were able to prove such a result for a special class of Φ in Lemma 4.3 (but see Remark 4.4). In general, we do not know the answer, but we can prove the following weaker result that suffices for us; the important feature is that the set E below contains the diagonal; hence we can make the equation $\Phi_{W'}(x_1, \dots, x_m) = \gamma$ hold (typically, at least) also when several, or all, x_i coincide.

Remark 5.2. The elimination of a null set in Problem 5.1 seems related to the infinite version of the (hypergraph) removal lemma [10], where the objective, in a different but related context, also is to replace a null set by an empty set.

Lemma 5.3. *Suppose that $\Phi\big((w_{ij})_{i<j}\big)$ is a multiaffine polynomial in the $\binom{m}{2}$ variables w_{ij}, $1 \le i < j \le m$, for some $m \ge 2$. Suppose further that $W : [0,1]^2 \to [0,1]$ is a graphon, i.e., a symmetric measurable function, and suppose that $\Phi_W(x_1, \dots, x_m) = \gamma$ for a.e. $x_1, \dots, x_m \in [0,1]$ and some $\gamma \in \mathbb{R}$. Then there is a version W' of W and a symmetric set $E \subseteq [0,1]^2$ such that $\lambda([0,1]^2 \setminus E) = 0$, $E \supseteq \{(x,x) : x \in [0,1]\}$, and $\Phi_{W'}(x_1, \dots, x_m) = \gamma$ for all $x_1, \dots, x_m$ such that $(x_i, x_j) \in E$ for every pair (i,j) with $1 \le i < j \le m$.*

The proof is rather technical, and is postponed until the Appendix.

As a consequence, we obtain a convenient criterion (patterned after [20]). We say that a graphon W is *finite-type*, or more specifically *k-type*, if there exists a partition of $[0, 1]$ into k sets $S_1, \dots, S_k$ such that W is constant on each rectangle $S_i \times S_j$. Making a rearrangement, we can without loss of generality assume that the sets S_i are intervals. (See [13] for a study of finite-type graph limits and the corresponding sequences of graphs, which generalise quasi-random graphs.)

Remark 5.4. In this paper, we consider for convenience only graphons defined on $[0, 1]$, but the definition extends to any probability space. Using this, we can equivalently, and more naturally, say that W is finite-type if it is equivalent to a graphon defined on a finite probability space.

Theorem 5.5. *Suppose that $\Phi\big((w_{ij})_{i<j}\big)$ is a multiaffine polynomial in the $\binom{m}{2}$ variables w_{ij}, $1 \le i < j \le m$, for some $m \ge 2$, and that $\gamma \in \mathbb{R}$. Then the following are equivalent.*

(i) *There exists a graphon W such that $\Phi_W(x_1, \dots, x_m) = \gamma$ for a.e. $x_1, \dots, x_m \in [0, 1]$, but W is not a.e. constant.*

(ii) *There exists a 2-type graphon W such that $\Phi_W(x_1, \dots, x_m) = \gamma$ for all $x_1, \dots, x_m$, but W is not (a.e.) constant.*

(iii) *There exist numbers $u, v, s \in [0, 1]$, not all equal, such that for every subset $A \subseteq [m]$, if we choose*

$$w_{ij} := \begin{cases} u, & i, j \in A, \\ v, & i, j \notin A, \\ s, & i \in A, j \notin A \text{ or conversely}, \end{cases} \tag{5.2}$$

then $\Phi((w_{ij})_{i<j}) = \gamma$.

In (ii), we may further require that the two parts of $[0, 1]$ are the intervals $[0, \frac{1}{2}]$ and $(\frac{1}{2}, 1]$.

The equivalence of (i) and (ii) shows that if a property of the type $\Phi_W = \gamma$ a.e. does not imply that W is a.e. constant (i.e., it is not a (mixed) quasi-random property for graphons), then there exists a counter-example that is a 2-type graphon. This generalises one of the results for induced subgraph counts by Simonovits and Sós [20].

Proof. (ii) $\iff$ (iii): A 2-type graphon W is defined by a partition (S_1, S_2) of $[0, 1]$ and three numbers $u, v, s \in [0, 1]$ such that $W = u$ on $S_1 \times S_1$, $W = v$ on $S_2 \times S_2$, and $W = s$ on $(S_1 \times S_2) \cup (S_2 \times S_1)$. It is easy to see that, for any S_1 and S_2 with $\lambda(S_1), \lambda(S_2) > 0$, such a graphon W satisfies $\Phi_W = \gamma$ if and only if $\Phi((w_{ij})_{i<j}) = \gamma$ with w_{ij} as in (5.2), for every choice of $A \subseteq [m]$. (Consider x_i such that $x_i \in S_1 \iff i \in A$.) Moreover, W is constant $\iff u = v = s$.
(ii) $\Rightarrow$ (i): trivial.
(i) $\Rightarrow$ (iii): suppose that W is a graphon as in (i) but that (iii) does not hold; we will show that this leads to a contradiction. Let W' and E be as in Lemma 5.3; for notational simplicity we replace W by W' and assume thus $W' = W$.

Suppose that $(x, y) \in E$. Given $A \subseteq [m]$, let $x_i := x$ for $i \in A$ and $x_i := y$ for $i \notin A$. Then $W(x_i, x_j) = w_{ij}$ as given by (5.2) with $u = W(x, x)$, $v = W(y, y)$, $s = W(x, y)$. Further, Lemma 5.3 shows that $\Phi((w_{ij})_{i<j}) = \Phi_W(x_1, \dots, x_m) = \gamma$. Since (iii) does not hold, no such u, v, s exist except with $u = v = s$. Consequently, we have shown the following property of W:

$$\text{If } (x, y) \in E, \text{ then } W(x, x) = W(y, y) = W(x, y). \tag{5.3}$$

Now suppose, more strongly, that (x_0, y_0) is a Lebesgue point of E, and that U is an open interval with $W(x_0, y_0) \in U$. It follows from the definition of Lebesgue points, that in a sufficiently small square Q centered at (x_0, y_0), the set $B := \{(x, y) \in Q : W(x, y) \in U\}$ has measure at least $\lambda(Q)/2$. Since $\lambda(E) = 1$, the same holds for $B \cap E$, and we may thus, by the regularity of the Lebesgue measure, find a compact set $K \subseteq B \cap E$ with $\lambda(K) > 0$. If $(x, y) \in K$, then $(x, y) \in E$, so by (5.3), $W(x, x) = W(x, y) \in U$. Consequently, if K' is the projection of K onto the first coordinate, then $W(x, x) \in U$ for $x \in K'$; furthermore, K' is a compact, and thus measurable, subset of $[0, 1]$, and $\lambda(K') > 0$.

By assumption, our W is not a.e. constant. Thus there exist two disjoint open intervals U_1 and U_2 such that $W^{-1}(U_\ell) := \{(x, y) : W(x, y) \in U_\ell\} \subseteq [0, 1]^2$ has positive measure, $\ell = 1, 2$. Then also, for each $\ell = 1, 2$, $D_\ell := E \cap W^{-1}(U_\ell)$ has positive measure, so we may pick a Lebesgue point (x_ℓ, y_ℓ) in D_ℓ. By what we just have shown, this implies that there exists a compact set $K_\ell \subseteq [0, 1]$ with $\lambda(K_\ell) > 0$ and $W(x, x) \in U_\ell$ for $x \in K_\ell$.

However, this means that if $(x, y) \in K_1 \times K_2$, then $W(x, x) \neq W(y, y)$, and thus by (5.3), $(x, y) \notin E$. Hence $E \cap (K_1 \times K_2) = \emptyset$. Since $\lambda(K_1 \times K_2) > 0$ and $\lambda(E) = 1$, this is a contradiction. □

Corollary 5.6. *Suppose that $\Phi\big((w_{ij})_{i<j}\big)$ is a multiaffine polynomial in the $\binom{m}{2}$ variables w_{ij}, $1 \le i < j \le m$, for some $m \ge 2$, and that $\gamma \in \mathbb{R}$. If every graphon W such that $\Phi_W(x_1, \dots, x_m) = \gamma$ for every $x_1, \dots, x_m \in [0, 1]$ is constant, then every graphon W such that $\Phi_W(x_1, \dots, x_m) = \gamma$ for a.e. $x_1, \dots, x_m \in [0, 1]$ is a.e. constant.*

In the terminology of Remark 3.9, if "$\Phi_W(x_1, \dots, x_m) = \gamma$ everywhere" is a (mixed) quasi-random property, then so is "$\Phi_W(x_1, \dots, x_m) = \gamma$ a.e. ". It is easily seen that the converse holds too; if W is a non-constant graphon such that $\Phi_W(x_1, \dots, x_m) = \gamma$ for every $x_1, \dots, x_m \in [0, 1]$, then there exists a non-constant m-type graphon with this property, and this graphon is not a.e. constant.

Proof. The assumption implies that there is no 2-type graphon W as in Theorem 5.5(ii), and thus there is no graphon W as in Theorem 5.5(i). □

5.1. *Further proofs of Theorem 3.2*

Instead of Lemma 4.3 we may use the weaker but more general Lemma 5.3; this lemma, with $\Phi((w_{ij})_{i<j}) := \prod_{ij\in E(F)} w_{ij}$, yields a version of W such that $\Psi_{F,W}(x_1, \dots, x_{|F|}) = p^{e(F)}$ at enough points so that the proof of Lemma 4.5 applies for a.e. (x, y). (Although Lemma 5.3 does not guarantee $\Psi_{F,W} = p^{e(F)}$ everywhere as Lemma 4.3 does.) This and Lemma 4.2 yield another proof of Theorem 3.2.

Alternatively, we may use Theorem 5.5 and argue as in the proof of Lemma 4.5, with only notational changes, to show that Theorem 5.5(iii) does not hold for this Φ, and hence by (i) $\iff$ (iii) in Theorem 5.5, W is a.e. constant and thus $W = p$ a.e., yielding another proof of Lemma 4.6, and thus of Theorem 3.2.

A modification of this argument is to use Lemma 4.5 as stated together with Corollary 5.6 to conclude that Lemma 4.6 holds.

Any of these proofs of Theorem 3.2 thus uses only the simple algebraic argument in Lemma 4.5 but combines it with results from this section. The latter results have rather long and technical proofs. If the objective is only to prove Theorem 3.2, the direct proof of Lemma 4.3 is much simpler than using Lemma 5.3 or one of its consequences Theorem 5.5 or Corollary 5.6. However, we have here started with the simplest case, and for other cases it seems much more complicated to prove analogues of Lemma 4.3 or Lemma 4.6 directly. Hence, our main method in what follows will be to use the results above, which once proven and available do not have to be modified.

Nevertheless, we have chosen to present also the direct proofs in Sections 4.2–4.4 in order to show alternative ways that in the case of Theorem 3.2 are simpler. We furthermore want to inspire readers to investigate whether there are similar direct proofs (that we have failed to find) in some of the cases treated later too.

6. One subset: proof of Theorem 3.1

We next give a proof of Theorem 3.1 along the lines of Section 4. We begin with a lemma giving an analogue of Lemma 4.1 for the case $A_1 = \cdots = A_m$.

If f is a function on $[0, 1]^m$ for some m, we let $\tilde{f}$ denote its symmetrisation defined by

$$\tilde{f}(x_1, \dots, x_m) := \frac{1}{m!} \sum_{\sigma\in\mathfrak{S}_m} f\big(x_{\sigma(1)}, \dots, x_{\sigma(m)}\big), \tag{6.1}$$

where $\mathfrak{S}_m$ is the symmetric group of all $m!$ permutations of $\{1, \dots, m\}$. Note that for any integrable f and any subset A of $[0, 1]$,

$$\int_{A^m} \tilde{f} = \int_{A^m} f. \tag{6.2}$$

Lemma 6.1. *Suppose that $f : [0, 1]^m \to \mathbb{R}$ is an integrable function such that $\int_{A^m} f = 0$ for all measurable subsets A of $[0, 1]$. Then $\tilde{f} = 0$ a.e.*

Proof. Let $A_1, \dots, A_m$ be disjoint subsets of $[0, 1]$. For any sequence $\xi_1, \dots, \xi_m \in \{0, 1\}^m$, take $A := \bigcup_{i:\xi_i=1} A_i$. Then $\mathbf{1}_A = \sum_{i=1}^m \xi_i \mathbf{1}_{A_i}$ and

$$0 = \int_{A^m} f = \int_{[0,1]^m} f \mathbf{1}_{A^m} = \sum_{i_1,\dots,i_m=1}^{m} \xi_{i_1} \dots \xi_{i_m} \int_{A_{i_1} \times \cdots \times A_{i_m}} f. \tag{6.3}$$

The monomials $\xi_{i_1} \dots \xi_{i_k}$ with $i_1 < \cdots < i_k$, $0 \le k \le m$, form a basis of the 2^m-dimensional space of functions on $\{0, 1\}^m$. Hence, collecting the terms in (6.3), the coefficient of each such monomial vanishes. In particular, for the coefficient of $\xi_1 \dots \xi_m$ we obtain a contribution only when $i_1, \dots, i_m$ is a permutation of $1, \dots, m$, and we obtain

$$0 = \sum_{\sigma \in \mathfrak{S}_m} \int_{A_{\sigma(1)} \times \cdots \times A_{\sigma(m)}} f = m! \int_{A_1 \times \cdots \times A_m} \tilde{f}.$$

The result follows by Lemma 4.1, applied to $\tilde{f}$. □

We can now translate property (3.1) to graphons, cf. Lemma 4.2.

Lemma 6.2. *Suppose that $G_n \to W$ for some graphon W and let F be a fixed graph and $\gamma \ge 0$ a fixed number. Then the following are equivalent:*

(i) *For all subsets U of $V(G_n)$,*

$$N(F, G_n; U) = \gamma |U|^{|F|} + o\big(|G_n|^{|F|}\big).$$

(ii) *For all subsets A of $[0, 1]$,*

$$\int_{A^{|F|}} \Psi_{F,W}(x_1, \dots, x_{|F|}) = \gamma \lambda(A)^{|F|}.$$

(iii) *$\widetilde{\Psi}_{F,W}(x_1, \dots, x_{|F|}) = \gamma$ for a.e. $x_1, \dots, x_{|F|} \in [0, 1]^{|F|}$, where $\widetilde{\Psi}_{F,W}$ is the symmetrisation of $\Psi_{F,W}$; see (6.1).*

Proof. This is proved almost exactly as in Lemma 4.2, with obvious notational changes and with Lemma 4.1 replaced by Lemma 6.1, which together with (6.2) implies (ii) $\iff$ (iii). The main difference is that we now use a single random set $B := \bigcup_{j:J_j=1} I''_{nj}$, where $\{J_j\}$ is a family of independent indicator variables. Hence, the analogue of (4.4) is not exact; we have

$$\mathbb{E} \prod_{i=1}^{|F|} J_{j_i} = \prod_{i=1}^{|F|} a_{j_i} \tag{6.4}$$

when $j_1, \dots, j_{|F|}$ are distinct, but in general not when two or more are equal. However, there are only $O(|G_n|^{|F|-1})$ choices of indices with at least two coinciding, and each such choice introduces an error that is at most $\lambda(I''_{nj_1} \times \cdots \times I''_{nj_{|F|}}) = |G_n|^{-|F|}$. Hence, we now have

$$\int_{A^n} (\Psi_{F,W_n} - \gamma) = \mathbb{E} \int_{B^n} (\Psi_{F,W_n} - \gamma) + o(1). \tag{6.5}$$

The error $o(1)$ is unimportant, and, assuming (i), the conclusion of (4.4) is valid in the form $\int_{A^n} (\Psi_{F,W_n} - \gamma) = o(1)$, which yields (ii) as in Section 4. □

We do not know any direct proof of the analogue of Lemma 4.3 for $\widetilde{\Psi}_{F,W}$. (This result follows by Lemma 6.2 and Theorem 3.1 once the latter is proven.) However, as in Section 5.1 we can nevertheless use the following lemma, which is a strengthening of Lemma 4.5.

Lemma 6.3. *Let F be a graph with $e(F) > 0$ and let W be a graphon. If $\widetilde{\Psi}_{F,W}(x_1, \dots, x_{|F|}) = p^{e(F)}$ for every $(x_1, \dots, x_{|F|}) \in [0, 1]^{|F|}$, then $W = p$.*

Proof. As in the proof of Lemma 4.5, first take $x_1 = \cdots = x_{|F|} = x$. Then $\widetilde{\Psi}_{F,W}(x_1, \dots, x_{|F|}) = \Psi_{F,W}(x_1, \dots, x_{|F|}) = W(x, x)^{e(F)}$, and thus $W(x, x) = p$. Using this, it is easy to see that if we take $x_1 = x$ and $x_2 = \cdots = x_{|F|} = y$, and d_i is the degree of vertex i, then

$$p^{e(F)} = \widetilde{\Psi}_{F,W}(x_1, \dots, x_{|F|}) = \frac{1}{|F|} \sum_{i \in V(F)} \left(\frac{W(x, y)}{p} \right)^{d_i} p^{e(F)}.$$

Since the right-hand side is a strictly increasing function of $W(x, y)$, this equation has only the solution $W(x, y) = p$. □

As in Section 4 there is a companion result where we allow exceptional null sets.

Lemma 6.4. *Let F be a graph with $e(F) > 0$ and let W be a graphon. If $\widetilde{\Psi}_{F,W}(x_1, \dots, x_{|F|}) = p^{e(F)}$ for a.e. $(x_1, \dots, x_{|F|}) \in [0, 1]^{|F|}$, then $W = p$ a.e.*

Proof. We have not tried to find a direct proof, since this follows directly from Lemma 6.3 and Corollary 5.6. □

Theorem 3.1 now follows from Lemmas 6.2 and 6.4. (Alternatively, we may use Lemma 5.3 or Theorem 5.5(iii) and argue as in the proof of Lemma 6.3.)

7. Further variations

7.1. Disjoint subsets

In Section 4 the sets $U_1, \dots, U_{|F|}$ of vertices were arbitrary and in Section 6 they were assumed to coincide. The opposite extreme is to require that they are disjoint. We can translate this version too to graphons as follows. Note that (iii) in the following lemma is the same as Lemma 4.2(iii); hence the two lemmas together show that it is equivalent to assume (4.1) (or (3.2)) for disjoint $U_1, \dots, U_{|F|}$ only; this implies the general case.

Lemma 7.1. *Suppose that $G_n \to W$ for some graphon W and let F be a fixed graph and $\gamma \geq 0$ a fixed number. Then the following are equivalent:*

(i) *For all disjoint subsets $U_1, \dots, U_{|F|}$ of $V(G_n)$,*

$$N(F, G_n; U_1, \dots, U_{|F|}) = \gamma \prod_{i=1}^{|F|} |U_i| + o\big(|G_n|^{|F|}\big).$$

(ii) *For all disjoint subsets $A_1, \dots, A_{|F|}$ of $[0, 1]$,*

$$\int_{A_1 \times \cdots \times A_{|F|}} \Psi_{F,W}(x_1, \dots, x_{|F|}) = \gamma \prod_{i=1}^{|F|} \lambda(A_i).$$

(iii) $\Psi_{F,W}(x_1, \dots, x_{|F|}) = \gamma$ *for a.e.* $x_1, \dots, x_{|F|} \in [0, 1]^{|F|}$.

Proof. Again we follow the proof of Lemma 4.2. The only difference is that we consider only disjoint sets $U_1, \dots, U_{|F|}$, etc. In particular, given disjoint subsets $A_1, \dots, A_{|F|}$ of $[0, 1]$, we want to construct the random sets B_i so that they too are disjoint. We do this by taking, for each j, the 0–1 random variables J_{ij} dependent, so that $\sum_i J_{ij} \leq 1$. (This is possible because $\sum_i a_{ij} \leq 1$ when $A_1, \dots, A_{|F|}$ are disjoint.) The vectors $(J_{ij})_{i=1}^{|F|}$ for different j are chosen independent as before. Just as in the proof of Lemma 6.2, the dependency among the J_{ij} means that (4.4) is not exact: in analogy with (6.4), $\mathbb{E} \prod_{i=1}^{|F|} J_{ij_i} = \prod_{i=1}^{|F|} a_{ij_i}$ when $j_1, \dots, j_{|F|}$ are distinct, but not in general. However, again as in the proof of Lemma 6.2, the total error is $o(1)$, so the analogue of (6.5) holds, and thus the conclusion $\int_{A_1\times\cdots\times A_{|F|}}(\Psi_{F,W_n}-\gamma)=o(1)$ of (4.4) holds for all disjoint sets $A_1,\ldots,A_{|F|}$.

Finally, for (ii) ⇒ (iii), note that Lemma 4.1 already is stated so that it suffices to consider disjoint $A_1,\ldots,A_{|F|}$. □

Lemma 7.1, combined with the remainder of the proof of Theorem 3.2 in Section 4, shows that in Theorem 3.2, it is sufficient to assume (3.2) for disjoint $U_1,\ldots,U_{|F|}$.

7.2. Sets of the same size

Another variation of Theorem 3.2 is to consider only subsets $U_1,\ldots,U_{|F|}$ of the same size. (We may combine this with the preceding variation and require that the sets are disjoint too.) This can be translated to considering only subsets $A_1,\ldots,A_{|F|}$ of the same measure by the same method as in the next subsection, when we further let the common size be a given number. Since we obtain stronger results in the next subsection, we leave the details to the reader.

7.3. Sets of a given size

Another variation of Theorem 3.2 is Theorem 3.6 where we consider only subsets $U_1,\ldots,U_{|F|}$ of a given size, which we assume is a fixed fraction α of $|G_n|$ (rounded to an integer). This is translated to graphons as follows.

Lemma 7.2. *Suppose that $G_n\to W$ for some graphon W and let F be a fixed graph and $\gamma\geq 0$ and $\alpha\in(0,1)$ be fixed numbers. Then the following are equivalent:*

(i) *For all subsets $U_1,\ldots,U_{|F|}$ of $V(G_n)$ with $|U_i|=\lfloor\alpha|G_n|\rfloor$,*

$$N(F,G_n;U_1,\ldots,U_{|F|})=\gamma\prod_{i=1}^{|F|}|U_i|+o\big(|G_n|^{|F|}\big). \tag{7.1}$$

(ii) *For all subsets $A_1,\ldots,A_{|F|}$ of $[0,1]$ with $\lambda(A_i)=\alpha$,*

$$\int_{A_1\times\cdots\times A_{|F|}}\Psi_{F,W}(x_1,\ldots,x_{|F|})=\gamma\prod_{i=1}^{|F|}\lambda(A_i). \tag{7.2}$$

(iii) *$\Psi_{F,W}(x_1,\ldots,x_{|F|})=\gamma$ for a.e. $x_1,\ldots,x_{|F|}\in[0,1]^{|F|}$.*

If $\alpha<1/|F|$, we may further, as in Lemma 7.1, in (i) *and* (ii) *add the requirement that the sets be disjoint.*

Proof. The equivalence (i) $\iff$ (ii) is proved as in the proof of Lemma 4.2, but some care has to be taken with the sizes and measures of the sets. We note that for any sets $A_1,\ldots,A_{|F|}$ and $A_1',\ldots,A_{|F|}'$,

$$\left|\int_{A_1\times\cdots\times A_{|F|}}\Psi_{F,W_n}-\int_{A_1'\times\cdots\times A_{|F|}'}\Psi_{F,W_n}\right|\leq\sum_{i=1}^{|F|}\lambda(A_i\mathbin{\triangle}A_i'). \tag{7.3}$$

Hence, we can modify the sets without affecting the results as long as the difference has measure $o(1)$. We argue as follows.

We obtain as in Section 4 that (i) is equivalent to (4.3), now for all subsets U_i'' of $[0,1]$ that are unions of sets I_{nj}'' and have measures $\lambda(U_i'')=\lfloor\alpha|G_n|\rfloor/|G_n|$. If (ii) holds, we may for any such U_i'' find $A_i\supseteq U_i''$ with $\lambda(A_i)=\alpha$; then (7.2) implies first (4.5) and then (4.3) by (7.3).

Conversely, given $A_1,\ldots,A_{|F|}$ with measures $\lambda(A_i)=\alpha$, the random sets B_i constructed above (either as in Section 4 or as in Section 7.1 in the disjoint case) have measures that are random but well concentrated:

$$\mathbb{E}\lambda(B_i) = \sum_j \mathbb{E}J_{ij}\lambda(I''_{nj}) = \sum_j a_{ij}\lambda(I''_{nj}) = \lambda(A_i) = \alpha$$

$$\operatorname{Var}\lambda(B_i) = \sum_j \operatorname{Var}(J_{ij})\lambda(I''_{nj})^2 \le |G_n|^{-1} \to 0.$$

Hence, if $\delta_n := |G_n|^{-1/3}$, say, then by Chebyshev's inequality

$$\mathbb{P}(|\lambda(B_i) - \alpha| > \delta_n) \le \delta_n^{-2}\operatorname{Var}(\lambda(B_i)) \le \delta_n \to 0.$$

If $|\lambda(B_i) - \alpha| \le \delta_n$ for all i, we adjust B_i to a set U''_i with $\lambda(U''_i) = \lfloor \alpha |G_n| \rfloor / |G_n|$ so that $\lambda(B_i \triangle U''_i) \le \delta_n + |G_n|^{-1} \le 2\delta_n$, and thus

$$\int_{B_1\times\cdots\times B_{|F|}} \Psi_{F,W_n} = \int_{U''_1\times\cdots\times U''_{|F|}} \Psi_{F,W_n} + O(\delta_n).$$

Consequently, if (4.3) holds, then $\int_{B_1\times\cdots\times B_{|F|}} \Psi_{F,W_n} = \gamma\alpha^{|F|} + O(\delta_n) + o(1)$ whenever $|\lambda(B_i) - \alpha| \le \delta_n$ for all i, and thus

$$\begin{aligned}\mathbb{E}\int_{B_1\times\cdots\times B_{|F|}} \Psi_{F,W_n} &= \gamma\alpha^{|F|} + O(\delta_n) + o(1) + O\Big(\sum_{i=1}^{|F|} \mathbb{P}\big(|\lambda(B_i) - \alpha| > \delta_n\big)\Big)\\ &= \gamma\alpha^{|F|} + o(1).\end{aligned}$$

Hence, (4.5) holds, for $A_1, \dots, A_{|F|}$ with measures $\lambda(A_i) = \alpha$, and thus (ii) holds by the argument in Section 4.

This proves (i) $\iff$ (ii); we may add the requirement that the sets be disjoint by the argument in the proof of Lemma 7.1.

To see that (ii) $\iff$ (iii), we use the following analysis lemma. (This seems to be less well known than Lemma 4.1; we guess that it is known, but we have been unable to find a reference.) □

Lemma 7.3. *Let $\alpha \in (0, 1)$. Suppose that $f : [0, 1]^m \to \mathbb{R}$ is an integrable function such that $\int_{A_1\times\cdots\times A_m} f = 0$ for all sequences $A_1, \dots, A_m$ of measurable subsets of $[0, 1]$ such that $\lambda(A_1) = \cdots = \lambda(A_m) = \alpha$. Then $f = 0$ a.e.*

Moreover, if $\alpha < m^{-1}$, it is enough to consider disjoint $A_1, \dots, A_m$.

Proof. For $f \in L^1([0, 1]^m)$ and $A_1, \dots, A_m \subseteq [0, 1]$, let

$$f(A_1, \dots, A_m) := \int_{A_1\times\cdots\times A_m} f,$$

and define further the functions

$$f_{A_1}(x_2, \dots, x_m) := \int_{A_1} f(x_1, x_2, \dots, x_m)\,\mathrm{d}x_1$$

and

$$f^{A_2,\dots,A_m}(x_1) := \int_{A_2\times\cdots\times A_m} f(x_1, x_2, \dots, x_m)\,\mathrm{d}x_2 \dots \mathrm{d}x_m.$$

By Fubini's theorem,

$$f(A_1, \dots, A_m) = f_{A_1}(A_2, \dots, A_m) = f^{A_2,\dots,A_m}(A_1). \tag{7.4}$$

We will derive the lemma from the following claims, which we will prove by induction in m.

Let B be a measurable subset of $[0, 1]$, let $0 < \alpha < 1$ and let f be an integrable function on B^m.

(i) *If $\alpha < \lambda(B)$ and $f(A_1, \dots, A_m) = 0$ for all $A_1, \dots, A_m \subset B$ with $\lambda(A_1) = \cdots = \lambda(A_m) = \alpha$, then $f(A_1, \dots, A_m) = 0$ for all $A_1, \dots, A_m \subseteq B$.*

(ii) *If $m\alpha < \lambda(B)$ and $f(A_1, \dots, A_m) = 0$ for all disjoint $A_1, \dots, A_m \subset B$ with $\lambda(A_1) = \cdots = \lambda(A_m) = \alpha$, then $f(A_1, \dots, A_m) = 0$ for all disjoint $A_1, \dots, A_m \subset B$ with $\lambda(A_1), \dots, \lambda(A_m) \leq \alpha$.*

Consider first the case $m = 1$, in which cases (i) and (ii) have the same hypotheses: $\alpha < \lambda(B)$ and $f(A) = 0$ if $\lambda(A) = \alpha$. Suppose that $A_1, A_2 \subset B$ with $\lambda(A_1) = \lambda(A_2) \leq \delta := \frac{1}{2}(\lambda(B) - \alpha)$. Then $\lambda(B \setminus (A_1 \cup A_2)) \geq \lambda(B) - 2\delta = \alpha$, and we may thus find a set $A_0 \subseteq B \setminus (A_1 \cup A_2)$ with $\lambda(A_0) = \alpha - \lambda(A_1)$. The assumption yields $f(A_1 \cup A_0) = 0 = f(A_2 \cup A_0)$, and thus

$$f(A_1) = -f(A_0) = f(A_2). \tag{7.5}$$

If $A \subset B$ is given with $\lambda(A) \leq \delta$ and $\lambda(A) = \alpha/N$ for some integer N, let $A_1 = A$ and choose further sets $A_2, \dots, A_N \subset B$ of the same measure α/N and with $A_1, \dots, A_N$ disjoint. By (7.5), then $f(A_k) = f(A_1) = f(A)$ for every $k \leq N$, and thus, by the assumption,

$$0 = f\Big(\bigcup_{k=1}^{N} A_k\Big) = \sum_{k=1}^{N} f(A_k) = Nf(A).$$

Consequently, $f(A) = 0$ for every $A \subset B$ with $\lambda(A) \leq \delta$ and $\lambda(A) = \alpha/N$. If x_0 is a density point of B (i.e., a point in B that is a Lebesgue point of $\mathbf{1}_B$), then there is a sequence $\varepsilon_n \to 0$ such that $\lambda(B \cap (x_0 - \varepsilon_n, x_0 + \varepsilon_n)) = \alpha/n$, and thus by what we have just shown, $\int_{x_0-\varepsilon_n}^{x_0+\varepsilon_n} f\mathbf{1}_B = f(B \cap (x_0 - \varepsilon_n, x_0 + \varepsilon_n)) = 0$ for every n. If further x_0 is a Lebesgue point of $f\mathbf{1}_B$, then this implies $f(x_0) = f(x_0)\mathbf{1}_B(x_0) = 0$. Since a.e. $x_0 \in B$ satisfies these conditions, $f = 0$ a.e. on B, which of course is equivalent to $f(A) = 0$ for every $A \subseteq B$. This proves both (i) and (ii) for $m = 1$.

For $m > 1$, we use, as already said, induction, and assume that the claims are true for smaller m. To prove (i), we fix $A_1 \subset B$ with $\lambda(A_1) = \alpha$, and see by (7.4) that f_{A_1} satisfies the assumptions of (i) on B^{m-1}. Thus, by the induction hypothesis, $f_{A_1}(A_2, \dots, A_m) = 0$ for all $A_2, \dots, A_m \subseteq B$. Fixing now instead such $A_2, \dots, A_m$, (7.4) shows that $f^{A_2,\dots,A_m}(A_1) = 0$ for all $\lambda(A_1) \subset B$ with $\lambda(A_1) = \alpha$, and thus by the case $m = 1$, $f^{A_2,\dots,A_m}(A_1) = 0$ for all $\lambda(A_1) \subset B$. By (7.4) again, this proves the induction hypothesis. Thus (i) is proved in general.

To prove (ii), we again fix A_1, and see by (7.4) that f_{A_1} satisfies the assumptions of (ii) on $(B \setminus A_1)^{m-1}$, noting that $(m-1)\alpha < \lambda(B \setminus A_1)$. Thus, by the induction hypothesis, $f_{A_1}(A_2, \dots, A_m) = 0$ for all disjoint $A_2, \dots, A_m \subseteq B \setminus A_1$ with $\lambda(A_k) \leq \alpha$ for every k. Hence, if we instead fix disjoint sets $A_2, \dots, A_m \subset B$ with $\lambda(A_k) \leq \alpha$ for every k, then (7.4) shows that $f^{A_2,\dots,A_m}(A_1) = 0$ for every $A_1 \subset B \setminus (A_2 \cup \cdots \cup A_m)$ with $\lambda(A_1) = \alpha$, and thus by the case $m = 1$, $f^{A_2,\dots,A_m}(A_1) = 0$ for every $A_1 \subset B \setminus (A_2 \cup \cdots \cup A_m)$ with $\lambda(A_1) \leq \alpha$. By (7.4) again, this proves the induction hypothesis, and (ii) is proved.

We have proved the claims above. We now take $B = [0, 1]$ and the lemma follows immediately by Lemma 4.1. □

Remark 7.4. When $\alpha = m^{-1}$, it is not enough to consider disjoint sets $A_1, \dots, A_m$ in Lemma 7.3. In fact, any f of the type $\sum_{i=1}^{m} g(x_i)$ where $\int_0^1 g = 0$ satisfies the assumption for such $A_1, \dots, A_m$. (We do not know whether these are the only possible f.) Taking W of this type and $F = K_2$, so that $\Psi_{F,W} = W$, we get a counter-example to Lemma 7.2, and to Theorem 3.6, for disjoint sets and $\alpha = 1/|F|$; see also Section 9 where this example reappears in a different formulation. We do not know whether there are such counter-examples for other graphs F.

Proof of Theorem 3.6. Theorem 3.6 follows by using Lemma 7.2 instead of Lemma 4.2 in (any version of) the proof of Theorem 3.2 in Section 4 (or 5.1). □

7.4. A single subset of a given size

The corresponding variation of Theorem 3.1 is Theorem 3.5 where we consider a single subset U with a given fraction α of the vertices. Again, there is a straightforward translation to graphons.

Lemma 7.5. *Suppose that $G_n \to W$ for some graphon W and let F be a fixed graph and $\gamma \geq 0$ and $\alpha \in (0,1)$ be fixed numbers. Then the following are equivalent:*

(i) *For every subset U of $V(G_n)$ with $|U| = \lfloor \alpha |G_n| \rfloor$,*

$$N(F, G_n; U) = \gamma |U|^{|F|} + o\big(|G_n|^{|F|}\big).$$

(ii) *For every subset A of $[0,1]$ with $\lambda(A) = \alpha$,*

$$\int_{A^{|F|}} \Psi_{F,W}(x_1, \dots, x_{|F|}) = \gamma \lambda(A)^{|F|}.$$

(iii) $\widetilde{\Psi}_{F,W}(x_1, \dots, x_{|F|}) = \gamma$ *for a.e.* $x_1, \dots, x_{|F|} \in [0,1]^{|F|}$.

Proof. The equivalence (i) $\iff$ (ii) is proved as for Lemma 7.2, using single sets U, A and B as in the proof of Lemma 6.2.

The equivalence (ii) $\iff$ (iii) follows by the following lemma, which strengthens Lemma 6.1 by considering subsets of a given size only. $\square$

Lemma 7.6. *Let $\alpha \in (0,1)$. Suppose that $f : [0,1]^m \to \mathbb{R}$ is an integrable function such that $\int_{A^m} f = 0$ for all measurable subsets A of $[0,1]$ with $\lambda(A) = \alpha$. Then $\tilde{f} = 0$ a.e.*

Proof. We begin by showing that the vanishing property extends to sets A with measure greater than α as follows:

$$\text{If } A \subseteq [0,1] \text{ with } \lambda(A) = r\alpha \text{ for some rational } r \geq 1, \text{ then } \int_{A^m} f = 0. \tag{7.6}$$

(The restriction to rational r may easily be removed by continuity, but it will suffice for us.) To see this, let N be an integer such that $M := rN$ is an integer, and partition A into M subsets $A_1, \dots, A_M$ of equal measure $\lambda(A_i) = \lambda(A)/M = r\alpha/M = \alpha/N$. Pick N of the sets A_i at random (uniformly over all $\binom{M}{N}$ possibilities), and let B be their union. Thus B is a random subset of $[0,1]$ with $\lambda(B) = \alpha$, and thus by the assumption $\int_{B^m} f = 0$. Taking the expectation we find

$$0 = \mathbb{E} \int_{B^m} f = \sum_{i_1, \dots, i_m = 1}^{M} \mathbb{P}(A_{i_1}, \dots, A_{i_m} \subseteq B) \int_{A_{i_1} \times \cdots \times A_{i_m}} f. \tag{7.7}$$

If $i_1, \dots, i_m$ are distinct, then, letting $(N)_m$ denote the falling factorial,

$$\mathbb{P}(A_{i_1}, \dots, A_{i_m} \subseteq B) = \frac{(N)_m}{(M)_m} = \left(\frac{N}{M}\right)^m + O\left(\frac{1}{N}\right) = r^{-m} + O\left(\frac{1}{N}\right).$$

This fails if two or more of $i_1, \dots, i_m$ coincide (in fact, the probability is $(N)_\nu/(M)_\nu \approx r^{-\nu}$, where ν is the number of distinct indices among $i_1, \dots, i_m$), so we let $U_N \subseteq [0,1]^m$ be the union of all $A_{i_1} \times \cdots \times A_{i_m}$ with at least two coinciding indices. By (7.7),

$$\begin{aligned}\frac{(N)_m}{(M)_m} \int_{A^m} f &= \sum_{i_1, \dots, i_m = 1}^{M} \frac{(N)_m}{(M)_m} \int_{A_{i_1} \times \cdots \times A_{i_m}} f \\ &= \sum_{i_1, \dots, i_m = 1}^{M} \left(\frac{(N)_m}{(M)_m} - \mathbb{P}(A_{i_1} \times \cdots \times A_{i_m} \subseteq B) \right) \int_{A_{i_1} \times \cdots \times A_{i_m}} f,\end{aligned}$$

and thus

$$\left| \frac{(N)_m}{(M)_m} \int_{A^m} f \right| \leq \int_{U_N} |f|. \tag{7.8}$$

Now let $N \to \infty$ (with rN integer). Note that $\lambda(U_N) \leq \binom{m}{2} N^{m-1}(\alpha/N)^m \leq \binom{m}{2}/N$. Thus $\lambda(U_N) \to 0$ and hence, since f is integrable, $\int_{U_N} |f| \to 0$. It follows from (7.8) and $(N)_m/(M)_m \to r^{-m}$ that $r^{-m}\int_{A^m} f = 0$, which proves (7.6).

Next, let $A_1, \ldots, A_m$ be arbitrary disjoint subsets of $[0, 1]$ with equal measure $\lambda(A_1) = \cdots = \lambda(A_m) = q\alpha$, for some rational q such that $(1+mq)\alpha \leq 1$. Choose $A_0 \subseteq [0, 1] \setminus \bigcup_1^m A_i$ with $\lambda(A_0) = \alpha$. For any sequence $\xi_1, \ldots, \xi_m \in \{0, 1\}^m$, let $\xi_0 := 1$ and take $A := \bigcup_{i\geq 0:\xi_i=1} A_i$. Then $\mathbf{1}_A = \sum_{i=0}^m \xi_i \mathbf{1}_{A_i}$ and we argue as in the proof of Lemma 6.1 with an extra set A_0: we have

$$0 = \int_{A^m} f = \int_{[0,1]^m} f\mathbf{1}_{A^m} = \sum_{i_1,\ldots,i_m=0}^{m} \xi_{i_1}\ldots\xi_{i_m}\int_{A_{i_1}\times\cdots\times A_{i_m}} f. \tag{7.9}$$

As in the proof of Lemma 6.1, it follows that the coefficient of $\xi_1\ldots\xi_m$ in (7.9) must vanish, and this coefficient comes from the terms where $i_1, \ldots, i_m$ is a permutation of $1, \ldots, m$. We thus obtain

$$0 = \sum_{\sigma\in\mathfrak{S}_m}\int_{A_{\sigma(1)}\times\cdots\times A_{\sigma(m)}} f = m!\int_{A_1\times\cdots\times A_m} \tilde{f}.$$

The result follows by Lemma 4.1 or 7.3, applied to $\tilde{f}$. □

Proof of Theorem 3.5. Theorem 3.5 follows by combining Lemmas 7.5 and 6.4, cf. Section 6. □

8. Induced subgraph counts

When considering counts of induced subgraphs, we translate the conditions to graphons similarly as above.

Lemma 8.1. *Suppose that $G_n \to W$ for some graphon W and let F be a fixed graph and $\gamma \geq 0$ a fixed number. Then the following are equivalent:*

(i) *For all subsets $U_1, \ldots, U_{|F|}$ of $V(G_n)$,*

$$N^*(F, G_n; U_1, \ldots, U_{|F|}) = \gamma\prod_{i=1}^{|F|}|U_i| + o\big(|G_n|^{|F|}\big).$$

(ii) *For all subsets $A_1, \ldots, A_{|F|}$ of $[0, 1]$,*

$$\int_{A_1\times\cdots\times A_{|F|}} \Psi^*_{F,W}(x_1, \ldots, x_{|F|}) = \gamma\prod_{i=1}^{|F|}\lambda(A_i).$$

(iii) *$\Psi^*_{F,W}(x_1, \ldots, x_{|F|}) = \gamma$ for a.e. $x_1, \ldots, x_{|F|} \in [0, 1]^{|F|}$.*

We may further in (i) *and* (ii) *add the conditions that, as in Lemma 7.1, the sets be disjoint, or that, as in Lemma 7.2, $|U_i| = \lfloor\alpha|G_n|\rfloor$ and $\lambda(A_i) = \alpha$ for a fixed $\alpha \in (0, 1)$, or, provided $\alpha < 1/|F|$, both.*

Proof. As for Lemma 4.2, using (2.9) instead of (2.6), and with the extra conditions treated as for Lemmas 7.1 and 7.2. □

Lemma 8.2. *Suppose that $G_n \to W$ for some graphon W and let F be a fixed graph and $\gamma \geq 0$ a fixed number. Then the following are equivalent:*

(i) *For all subsets U of $V(G_n)$,*

$$N^*(F, G_n; U) = \gamma|U|^{|F|} + o\big(|G_n|^{|F|}\big).$$

(ii) *For all subsets A of $[0, 1]$,*

$$\int_{A^{|F|}} \Psi^*_{F,W}(x_1, \ldots, x_{|F|}) = \gamma\lambda(A)^{|F|}.$$

(iii) *$\widetilde{\Psi}^*_{F,W}(x_1, \ldots, x_{|F|}) = \gamma$ for a.e. $x_1, \ldots, x_{|F|} \in [0, 1]^{|F|}$.*

We may further in (i) *and* (ii) *add the conditions that, as in Lemma 7.5, $|U| = \lfloor \alpha |G_n| \rfloor$ and $\lambda(A) = \alpha$ for a fixed $\alpha \in (0, 1)$.*

Proof. As for Lemma 6.2, using (2.9) instead of (2.6), and with the extra size conditions treated as for Lemma 7.5, using Lemma 7.6. □

However, it is now more complicated to do the algebraic step, i.e., to solve the equations in (iii) in these lemmas; the reason is that $\Psi^*_{F,W}$ and $\widetilde{\Psi}^*_{F,W}$ are not monotone in W. For $\Psi^*_{F,W}$, we can argue as follows. (See also the somewhat different argument in [17].)

Lemma 8.3. *Let F be a graph with $|F| > 1$, let W be a graphon and let $p \in (0, 1)$. If $\Psi^*_{F,W}(x_1, \ldots, x_{|F|}) = p^{e(F)}(1-p)^{\binom{|F|}{2}-e(F)}$ for every $x_1, \ldots, x_{|F|} \in [0, 1]^{|F|}$, then either $W = p$ or $W = \bar{p}$.*

Proof. First, take all x_i equal. Recalling the definitions (2.7) and (3.3), we see that

$$\Psi^*_{F,W}(x, \ldots, x) = W(x, x)^{e(F)}(1 - W(x, x))^{e(\overline{F})} = \beta_F(W(x, x)).$$

Thus, $\beta_F(W(x, x)) = \beta_F(p)$, and hence, cf. the definition of the conjugate $\bar{p}$ in Section 3, $W(x, x) \in \{p, \bar{p}\}$ for every x.

Next, if vertex i has degree d_i and we choose $x_i = y$ and $x_j = x$ for $j \neq i$, then

$$\Psi^*_{F,W}(x_1, \ldots, x_{|F|}) = \left(\frac{W(x, y)}{W(x, x)}\right)^{d_i} \left(\frac{1 - W(x, y)}{1 - W(x, x)}\right)^{|F|-1-d_i} \Psi^*_{F,W}(x, \ldots, x),$$

and thus

$$\left(\frac{W(x, y)}{W(x, x)}\right)^{d_i} \left(\frac{1 - W(x, y)}{1 - W(x, x)}\right)^{|F|-1-d_i} = 1, \quad i \in V(F). \tag{8.1}$$

If F is not regular, we may choose vertices i and j with $d_i \neq d_j$. Taking logarithms of (8.1) and the same equation with i replaced by j, we obtain a non-singular homogeneous system of linear equations in $\log(W(x, y)/W(x, x))$ and $\log((1 - W(x, y))/(1 - W(x, x)))$, and thus these logarithms vanish, so $W(x, y) = W(x, x)$ for every x and y in $[0, 1]$. Hence, if $x, y \in [0, 1]$, then $W(x, x) = W(x, y) = W(y, x) = W(y, y)$, and it follows that W is constant, and thus either $W = p$ or $W = \bar{p}$.

It remains to treat the case when F is regular, $d_i = d$ for all i. Note first that if F is a complete graph, then $\Psi^*_{F,W} = \Psi_{F,W}$, and the result follows by Lemma 4.5. Further, if F is empty, the result follows by taking complements, replacing F by $\overline{F}$, which is complete, W by $1 - W$, and p by $1 - p$. We may thus assume that $1 \le d \le |F| - 2$.

We now choose two vertices $i, j \in V(F)$ and let $x_i = x_j = y$ and $x_k = x$, $k \neq i, j$. If there is an edge $ij \in E(F)$, then

$$\Psi^*_{F,W}(x_1, \ldots, x_{|F|}) = \left(\frac{W(x, y)}{W(x, x)}\right)^{2d-2} \left(\frac{1 - W(x, y)}{1 - W(x, x)}\right)^{2(|F|-1-d)} \left(\frac{W(y, y)}{W(x, x)}\right) \Psi^*_{F,W}(x, \ldots, x),$$

and thus, using (8.1),

$$\frac{W(y, y)}{W(x, x)} = \left(\frac{W(x, y)}{W(x, x)}\right)^2$$

or

$$W(x, x)W(y, y) = W(x, y)^2. \tag{8.2}$$

Choosing instead $i, j \in V(F)$ with $ij \notin E(F)$, we similarly obtain

$$(1 - W(x, x))(1 - W(y, y)) = (1 - W(x, y))^2. \tag{8.3}$$

Subtracting (8.3) from (8.2) we find

$$W(x, x) + W(y, y) = 2W(x, y)$$

and thus, also using (8.2) again,

$$\begin{aligned}\big(W(x,x)-W(y,y)\big)^2 &= \big(W(x,x)+W(y,y)\big)^2 - 4W(x,x)W(y,y)\\ &= 4W(x,y)^2 - 4W(x,y)^2 = 0.\end{aligned}$$

Hence $W(x,x) = W(y,y)$ for all $x, y \in [0,1]$, which by (8.2) implies that W is a constant, which must be p or $\bar{p}$. □

As above, the results in the Appendix imply that we can relax the assumption to hold only almost everywhere.

Lemma 8.4. *Let F be a graph with $|F| > 1$, let W be a graphon and let $p \in (0,1)$. If $\Psi^*_{F,W}(x_1,\dots,x_{|F|}) = p^{e(F)}(1-p)^{\binom{|F|}{2}-e(F)}$ for a.e. $x_1,\dots,x_{|F|} \in [0,1]^{|F|}$, then either $W = p$ a.e. or $W = \bar{p}$ a.e.*

Proof. By Corollary 5.6 and Lemma 8.3, W has to be a constant c a.e. Then $\Psi^*_{F,W}(x_1,\dots,x_{|F|}) = \beta_F(c)$ a.e., and thus $\beta_F(c) = \beta_F(p)$; hence $c = p$ or $c = \bar{p}$. □

Proof of Theorems 3.11 and 3.12. As in Section 4, we may assume that $G_n \to W$ for some graphon W. By the assumption and Lemma 8.1, then

$$\Psi^*_{F,W}(x_1,\dots,x_{|F|}) = \beta_F(p) := p^{e(F)}(1-p)^{\binom{|F|}{2}-e(F)}$$

for a.e. $x_1,\dots,x_{|F|} \in [0,1]^{|F|}$, which by Lemma 8.4 implies either $W = p$ a.e. or $W = \bar{p}$ a.e. □

For $\widetilde{\Psi}^*_{F,W}$, the situation is even more complicated. In fact, Simonovits and Sós [20] showed that the path $P_3 = K_{1,2}$ and its complement $\overline{P}_3$ are not HI (recall Definition 3.7). Thus, the analogue of Lemma 8.3 for $\widetilde{\Psi}^*_{F,W}$ cannot hold in general.

We can, however, easily obtain the partial results of [20] by our methods. We note that by Theorem 5.5, it suffices to study 2-type graphons; equivalently, it suffices to study $\widetilde{\Psi}^*_{F,W}(x_1,\dots,x_{|F|})$ for sequences $x_1,\dots,x_{|F|}$ with at most two distinct values. For any sequence $x_1,\dots,x_{|F|}$ with $x_i = x$ for k values of i, and $x_i = y$ for the $|F|-k$ remaining values, we have

$$\widetilde{\Psi}^*_{F,W}(x_1,\dots,x_{|F|}) = \binom{|F|}{k}^{-1} Q_k\big(W(x,x), W(y,y), W(x,y)\big), \tag{8.4}$$

where $Q_k(u,v,s)$ is the polynomial, defined for a given graph F and $k = 0,\dots,|F|$, by

$$Q_k(u,v,s) = \sum_{\substack{A\subseteq V(F)\\ |A|=k}} u^{e(A)}(1-u)^{\binom{k}{2}-e(A)} v^{e(\bar{A})}(1-v)^{\binom{|F|-k}{2}-e(\bar{A})} s^{e(A,\bar{A})}(1-s)^{k(|F|-k)-e(A,\bar{A})},$$

where $\bar{A} := V(F)\setminus A$, $e(A)$ is the number of edges with both endpoints in A, and $e(A,\bar{A})$ is the number of edges with one endpoint in A and one in $\bar{A}$.

By symmetry, $Q_{|F|-k}(u,v,s) = Q_k(v,u,s)$. Note that $Q_0(u,v,s) = \beta_F(v)$ and $Q_{|F|}(u,v,s) = \beta_F(u)$. In particular, $Q_0(u,v,s) = \beta_F(p) \iff v \in \{p,\bar{p}\}$ and $Q_{|F|}(u,v,s) = \beta_F(p) \iff u \in \{p,\bar{p}\}$.

Remark 8.5. These polynomials are essentially the same as the polynomials $\mathbb{P}^k_{u,v}(s)$ defined by Simonovits and Sós [20]. More precisely,

$$\mathbb{P}^k_{u,v}(s) := \binom{|F|}{k} u^{e(F)}(1-u)^{e(\bar{F})} - Q_k(u,v,s).$$

Hence, the condition in Theorem 8.6(iv) below is equivalent to $\mathbb{P}^k_{u,v}(s) = 0$, with $u, v \in \{p,\bar{p}\}$.

Theorem 8.6. *Let F be a graph with $|F| > 1$ and let $0 < p < 1$. Then the following are equivalent:*

(i) *F is HI(p).*

(ii) *If $\Psi^*_{F,W}(x_1,\dots,x_{|F|}) = \beta_F(p)$ for a.e. $x_1,\dots,x_{|F|} \in [0,1]^{|F|}$, then either $W = p$ a.e. or $W = \bar{p}$ a.e.*

(iii) *If* $\Psi^*_{F,W}(x_1,\ldots,x_{|F|}) = \beta_F(p)$ *for all* $x_1,\ldots,x_{|F|} \in [0,1]^{|F|}$, *then either* $W = p$ *or* $W = \bar{p}$

(iv) *If* $Q_k(u,v,s) = \binom{|F|}{k}\beta_F(p)$ *for* $k = 1,\ldots,|F|-1$, *and* $u, v \in \{p, \bar{p}\}$, *then* $u = v = s$.

Proof. (i) $\iff$ (ii) follows by Lemma 8.2 and our general method.
(ii) $\iff$ (iii) follows by Corollary 5.6 (and the comment after it).
(ii) $\iff$ (iv) follows by Theorem 5.5, together with the remarks on Q_0 and $Q_{|F|}$ above. □

Proof of Theorem 3.13. Again we may assume that $G_n \to W$. It then follows by Lemma 8.2(i) $\iff$ (iii) and Theorem 8.6(i) ⇒ (ii) that either $W = p$ a.e. or $W = \bar{p}$ a.e. □

For $F = P_3$, it suffices by symmetry to check Q_1 in (iv); we find $Q_1(u,v,s) = 2vs(1-s) + (1-v)s^2$, and it is easy to find solutions with $u = v = p \neq s$; see [20] for details. On the other hand, Simonovits and Sós [20] have shown that every regular graph (and a few others) satisfies (iv), and thus is $\mathsf{HI}(p)$.

The algebraic problem of determining whether there are any other cases where the overdetermined system in Theorem 8.6(iv) has a non-trivial root is still unsolved.

9. Cuts

Chung and Graham [6] considered also $e_G(U,\overline{U})$, the number of edges in the graph G across a cut $(U,\overline{U})$, where $\overline{U} := V(G) \setminus U$. They proved the following results:

Theorem 9.1 (*Chung and Graham [6]*). *Suppose that* (G_n) *is a sequence of graphs with* $|G_n| \to \infty$ *and let* $0 \le p \le 1$. *Then* (G_n) *is p-quasi-random if and only if, for all subsets U of* $V(G_n)$,

$$e_{G_n}(U,\overline{U}) = p|U||\overline{U}| + o\big(|G_n|^2\big). \tag{9.1}$$

Theorem 9.2 (*Chung and Graham [6]*). *Let* $\alpha \in (0,1)$ *with* $\alpha \neq 1/2$. *Suppose that* (G_n) *is a sequence of graphs with* $|G_n| \to \infty$ *and let* $0 \le p \le 1$. *Then* (G_n) *is p-quasi-random if and only if* (9.1) *holds for all subsets U of* $V(G_n)$ *with* $|U| = \lfloor \alpha |G_n| \rfloor$.

However, as shown in [7,6], Theorem 9.2 does *not* hold for $\alpha = 1/2$.

Note that in our notation,

$$e_G(U,\overline{U}) = N(K_2, G; U, \overline{U}), \tag{9.2}$$

so these results are closely connected to Theorem 3.2 and its variants. We may use the methods above to show these results too, and to see why $\alpha = 1/2$ is an exception in Theorem 9.2.

We thus assume that $G_n \to W$ for some graphon W, and translate the properties above to properties of W. We state this as a lemma in the same style as earlier, and note that Theorems 9.1 and 9.2 are immediate consequences.

Lemma 9.3. *Suppose that* $G_n \to W$ *for some graphon W and let* $p \in [0,1]$. *Then the following are equivalent:*

(i) *For all subsets U of* $V(G_n)$,

$$e_{G_n}(U,\overline{U}) = p|U||\overline{U}| + o\big(|G_n|^2\big).$$

(ii) *For all subsets A of* $[0,1]$,

$$\int_{A\times\bar{A}} W(x,y) = p\lambda(A)\lambda(\bar{A}). \tag{9.3}$$

(iii) $W = p$ *a.e.*

For any fixed $\alpha \in (0,1) \setminus \{\frac{1}{2}\}$, *we may further add the condition that* $|U| = \lfloor \alpha |G_n| \rfloor$ *in* (i) *and* $\lambda(A) = \alpha$ *in* (ii). *(If we add these conditions with* $\alpha = 1/2$, *the equivalence* (i) $\iff$ (ii) *still holds, but these do not imply* (iii).*)*

Proof. The equivalence (i) $\iff$ (ii) follows as in Lemmas 4.2 and 7.1, arguing as in Lemma 7.2 in the case of a fixed size $\alpha \in (0, 1)$.

The implication (iii) $\Rightarrow$ (ii) is trivial, and (ii) $\Rightarrow$ (iii) follows by the following lemma, applied to $W - p$. $\square$

Lemma 9.4. *Let $\alpha \in (0, 1) \setminus \{\frac{1}{2}\}$. If $f : [0, 1]^2 \to \mathbb{R}$ is a symmetric measurable function such that $\int_{A\times([0,1]\setminus A)} f = 0$ for every subset A of $[0, 1]$ with $\lambda(A) = \alpha$, then $f = 0$ a.e.*

Proof. Let $f_1(x) := \int_0^1 f(x, y)\,\mathrm{d}y$ be the marginal of f. Then

$$0 = \int_{A\times([0,1]\setminus A)} f = \int_A f_1(x)\,\mathrm{d}x - \int_{A\times A} f(x, y)\,\mathrm{d}x\,\mathrm{d}y$$

$$= \int_{A\times A} \Big(\frac{1}{\alpha} f_1(x) - f(x, y)\Big)\,\mathrm{d}x\,\mathrm{d}y. \tag{9.4}$$

Lemma 7.6 now shows that the symmetrisation $\frac{1}{2\alpha} f_1(x) + \frac{1}{2\alpha} f_1(y) - f(x, y) = 0$ a.e., i.e.,

$$f(x, y) = \frac{1}{2\alpha}\big(f_1(x) + f_1(y)\big). \tag{9.5}$$

Integrating (9.5) with respect to both variables we find $\int f = \frac{2}{2\alpha} \int f$, and thus, because $\alpha < 1$, $\int f = 0$. Integrating (9.5) with respect to one variable we then find $f_1(x) = \frac{1}{2\alpha} f_1(x)$ a.e., and thus $f_1(x) = 0$ a.e. because $\alpha \neq 1/2$. A final appeal to (9.5) yields $f(x, y) = 0$ a.e. $\square$

This proof also shows what goes wrong with Theorem 9.2 when $\alpha = 1/2$. In this case, the condition of Lemma 9.4 still implies (9.5), but this is satisfied if (and only if) $f(x, y) = g(x) + g(y)$ for any integrable g with $\int g = 0$, and as a result we see that (9.1) is satisfied for all U with $|U| = \lfloor |G_n|/2 \rfloor$ whenever $G_n \to W$ where W is a graphon of the form $W(x, y) = h(x) + h(y)$ with $\int_0^1 h = p/2$. (One such example of (G_n), with $p = 1/2$ and $h(x) = \frac{1}{2}\mathbf{1}[x \geq 1/2]$ is given in [7,6].) Cf. Remark 7.4.

Remark 9.5. The condition that f is symmetric is essential in Lemma 9.4. If f is anti-symmetric, then (9.4) implies that f satisfies the condition if and only if $\int_0^1 f(x, y)\,\mathrm{d}y = 0$ for a.e. x. One example is $\sin(2\pi(x - y))$.

Chung et al. [7] remarked that Theorem 9.2 holds in the case $\alpha = 1/2$ too, if we further assume that (G_n) is almost regular (see below for definition). We discuss and show this in the next section.

10. The degree distribution

If G is a graph, let D_G denote the random variable defined as the degree d_v of a randomly chosen vertex v (with the uniform distribution on $V(G)$). Thus $0 \leq D_G \leq |G| - 1$, and we normalise D_G by considering $D_G/|G|$, which is a random variable in $[0, 1]$. If (G_n) is a sequence of graphs, with $|G_n| \to \infty$ as usual, we say that (G_n) has *asymptotic (normalised) degree distribution* μ if D_G tends to μ in distribution. (Here μ is a distribution, i.e., a probability measure, on $[0, 1]$.) In the special case when μ is concentrated at a point $p \in [0, 1]$, we say that (G_n) is *almost p-regular* (or *almost regular* if we do not want to specify p); this thus is the case if and only if $D_{G_n} \xrightarrow{\mathrm{p}} p$, with convergence in probability, which means that all but $o(|G_n|)$ vertices in G_n have degrees $p|G_n| + o(|G_n|)$. Since the random variables D_{G_n} are uniformly bounded (by 1), this is further equivalent to convergence in mean, and thus a sequence (G_n) is almost p-regular if and only if $\mathbb{E}|D_{G_n} - p| \to 0$, or, more explicitly, cf. [7],

$$\sum_{v\in V(G)} \big|d_v - p|G_n|\big| = o(|G_n|^2). \tag{10.1}$$

The normalised degree distribution behaves continuously under graph limits, and a corresponding "normalised degree distribution" may be defined for every graph limit too. (See further [9].) For a graphon W we define the marginal $w(x) := \int_0^1 W(x, y)\,dy$ and the random variable $D_W := w(U) = \int_0^1 W(U, y)\,dy$, where $U \sim U[0, 1]$ is uniformly distributed on $[0, 1]$.

Theorem 10.1. *If G_n are graphs with $|G_n| \to \infty$ and $G_n \to W$ for some graphon W, then $D_{G_n}/|G_n| \xrightarrow{d} D_W$. Hence, (G_n) has an asymptotical degree distribution, and this equals the distribution of the random variable $D_W := \int_0^1 W(U, y)\,dy$.*

Proof. It is easily seen that, for every $k \geq 1$, the moment $\mathbb{E}(D_G/|G|)^k$ equals $t(S_k, G)$, where $S_k = K_{1,k}$ is a star with $k+1$ vertices, and similarly the moment $\mathbb{E}W_G^k = t(S_k, W)$. Consequently, $\mathbb{E}(D_{G_n}/|G_n|)^k = t(S_k, G_n) \to t(S_k, W) = \mathbb{E}D_W$ for every $k \geq 1$, and thus $D_{G_n} \xrightarrow{d} D_W$ by the method of moments. □

Corollary 10.2. *Let (G_n) be a sequence of graphs and W a graphon such that $G_n \to W$. Then G_n is almost p-regular if and only if $\int_0^1 W(x, y)\,dy = p$ for a.e. $x \in [0, 1]$.* □

In particular, a quasi-random sequence of graphs is almost regular, but the converse does not hold.

Motivated by Corollary 10.2, we say that a graphon W is *p-regular* if its marginal $\int_0^1 W(x, y)\,dy = p$ a.e. This is evidently not a quasi-random property of graphons, but it can be used in conjunction with the failed case $\alpha = 1/2$ in Section 9. We find the following lemmas.

Lemma 10.3. *Let $\alpha \in (0, 1)$. If $f : [0, 1]^2 \to \mathbb{R}$ is a symmetric measurable function such that $\int_{A\times([0,1]\setminus A)} f = 0$ for every subset A of $[0, 1]$ with $\lambda(A) = \alpha$, and $\int_0^1 f(x, y)\,dy = 0$ for a.e. x, then $f = 0$ a.e.*

Proof. The proof of Lemma 9.4 shows that (9.5) holds, where now by assumption $f_1 = 0$. □

Lemma 10.4. *Let $p \in [0, 1]$ and $\alpha \in (0, 1)$. Suppose that (G_n) is an almost p-regular sequence of graphs and that $G_n \to W$ for some graphon W. Then the following are equivalent:*

(i) *For all subsets U of $V(G_n)$ with $|U| = \lfloor \alpha |G_n| \rfloor$,*

$$e_{G_n}(U, \overline{U}) = p\alpha(1-\alpha)|G_n|^2 + o\big(|G_n|^2\big). \tag{10.2}$$

(ii) *For all subsets A of $[0, 1]$ with $\lambda(A) = \alpha$,*

$$\int_{A\times\overline{A}} W(x, y) = p\alpha(1-\alpha).$$

(iii) *$W = p$ a.e.*

Proof. By Lemma 9.3, it remains only to show that (i) ⇒ (ii) in the case $\alpha = 1/2$. However, by Corollary 10.2, W is p-regular, so (ii) ⇒ (iii) follows by Lemma 10.3 applied to $W - p$. □

Lemma 10.4 yields, by our general machinery, immediately the following theorem by Chung et al. [7], which supplements Theorem 9.2 in the case $\alpha = 1/2$ (and otherwise is a trivial consequence of Theorem 9.2).

Theorem 10.5 (*Chung et al. [7]*)**.** *Let $0 \leq p \leq 1$ and $\alpha \in (0, 1)$. Suppose that (G_n) is a sequence of graphs with $|G_n| \to \infty$. Then (G_n) is p-quasi-random if and only if (G_n) is almost p-regular and (10.2) holds for all subsets U of $V(G_n)$ with $|U| = \lfloor \alpha |G_n| \rfloor$.*

Acknowledgements

This work was begun during the workshop on Graph Limits, Homomorphisms and Structures in Hraniční Zámeček, Czech Republic, 2009. We thank Asaf Shapira, Miki Simonovits, Vera Sós and Balázs Szegedy for interesting discussions.

Research partly done at Institut Mittag-Leffler, Djursholm, Sweden.

Appendix. Proof of Lemma 5.3

We first prove the following lemma, which is a (weak) substitute for the Lebesgue differentiation theorem when we consider points on the diagonal only. (The Lebesgue differentiation theorem says nothing about such points, since the diagonal is a null set. A simple counter-example is $W(x, y) = \mathbf{1}[x < y]$.) We introduce some further notation.

If $A \subseteq [0, 1]$ with $\lambda(A) > 0$, let λ_A be the normalised Lebesgue measure on A given by $\lambda_A(B) := \lambda(A \cap B)/\lambda(A)$, $B \subseteq [0, 1]$. (In other words, λ_A is the distribution of a uniform random point in A.)

The definition (2.10) of the cut norm generalises to arbitrary measure spaces. In particular, if $A \subseteq [0, 1]$ with $\lambda(A) > 0$, we let $\|W\|_{\square,A}$ denote the cut norm on $A \times A$ with respect to the normalised measure λ_A. More generally, if A and $B \subseteq [0, 1]$ have positive measures, then

$$\|W\|_{\square,A\times B} := \sup_{S\subseteq A,\, T\subseteq B} \int_{S\times T} W(x, y)\, \mathrm{d}\lambda_A(x)\, \mathrm{d}\lambda_B(y)$$

denotes the (normalised) cut norm on $A \times B$.

Lemma A.1. *For every $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon) > 0$ such that if $W : [0, 1]^2 \to [0, 1]$ is a symmetric and measurable function and $A \subseteq [0, 1]$ with $\lambda(A) > 0$, then there exists $B \subseteq A$ with $\lambda(B) \geq \delta\lambda(A)$ and a real number $w \in [0, 1]$ such that $\|W - w\|_{\square,B} < \varepsilon$.*

Remark A.2. The example $W(x, y) = \mathbf{1}[x < y]$ shows that Lemma A.1 in general fails for non-symmetric functions.

Remark A.3. Lemma A.1 is not true with the stronger conclusion obtained by replacing cut norm by L^1 norm. An example is (whp) given, for any $\varepsilon < 1/2$, by the 0/1-valued function W corresponding to a random graph $G(n, 1/2)$, for a large n.

Although Lemma A.1 is a purely analytic statement, we prove it by using combinatorial methods; in fact, the proof is an adaption of the relevant parts of the proof of one of the main theorems in Simonovits and Sós [20] to graphons (instead of graphs).

Proof. By considering the restriction of W to $A \times A$ and a measure preserving bijection of (A, λ_A) onto $([0, 1], \lambda)$, it suffices to consider the case $A = [0, 1]$.

Let $r = \lceil 3/\varepsilon \rceil$ and let M be the Ramsey number $R(r; r) = R(r, \dots, r)$ (with r repeated r times); in other words, every colouring of the edges of the complete graph K_M with at most r colours contains a monochromatic K_r. (See e.g. [12].)

By the (strong) analytic Szemerédi regularity lemma by Lovász and Szemerédi [15, Lemma 3.2], there is an integer $K = k(\varepsilon/(4M^2))$ (depending on ε only, since M is a function of ε) and, for some $k \leq K$, a partition $\mathcal{P} = \{S_1, \dots, S_k\}$ of $[0, 1]$ into k sets of equal measure $1/k$ with the property that for every set $R \subseteq [0, 1]^2$ that is a union of at most k^2 rectangles, we have

$$\left| \int_R (W - W_{\mathcal{P}}) \right| \leq \frac{\varepsilon}{4M^2}, \tag{A.1}$$

where $W_{\mathcal{P}}$ is the function that is constant on each set $S_i \times S_j$ and equal to the average $k^2 \int_{S_i\times S_j} W$ there. (That is, $W_{\mathcal{P}}$ is the conditional expectation of W given the σ-field generated by $\{S_i \times S_j\}_{i,j=1}^k$.) Let w_{ij} be this average $k^2 \int_{S_i\times S_j} W$. We consider two cases separately:

(i): $k \geq 2M$. Let, for $i, j = 1, \dots, k$,

$$d_{ij} := \|W - W_{\mathcal{P}}\|_{\square,S_i\times S_j} = \|W - w_{ij}\|_{\square,S_i\times S_j} = \max\bigl(d_{ij}^+, d_{ij}^-\bigr), \tag{A.2}$$

where

$$d_{ij}^{\pm} := \sup_{S\subseteq S_i,\, T\subseteq S_j} \pm k^2 \int_{S\times T} (W - W_{\mathcal{P}}).$$

It follows from (A.1) that

$$\sum_{i,j=1}^{k} d_{ij}^{+} \le k^2 \frac{\varepsilon}{4M^2},$$

and thus the number of pairs (i, j) with $d_{ij}^{+} > \varepsilon/3$ is less than k^2/M^2, and similarly for d_{ij}^{-}.

Say that a pair (i, j) is *bad* if $d_{ij} > \varepsilon/3$ or $i = j$, and *good* otherwise. By (A.2), the number of bad pairs is thus less than $2k^2/M^2 + k \le k^2/M$, using our assumption that $k \ge 2M$ and assuming, as we may, that $M \ge 4$.

Consider the graph H on $[k]$ where there is an edge ij whenever (i, j) is a good pair. Further, give every edge ij in H the colour $c_{ij} := \max(\lceil r w_{ij} \rceil, 1) \in [r]$. Since H has more than $\frac{1}{2}(k^2 - \frac{1}{M}k^2) = (1 - \frac{1}{M})\frac{k^2}{2}$ edges, Turán's theorem shows that H contains a complete subgraph K_M, and the choice of M implies that this complete subgraph contains a complete monochromatic subgraph K_r.

In other words, there is a $c \in [r]$ such that, after renumbering the sets S_i in $\mathcal{P}$, for all $i, j \in [r]$ with $i \ne j$, (i, j) is a good pair and $c_{ij} = c$. Let $w := c/r \in [0, 1]$. Then, for $1 \le i < j \le r$, $c - 1 \le r w_{ij} \le c$, so $|w_{ij} - w| \le 1/r \le \varepsilon/3$. Since (i, j) is good, this further implies

$$\|W - w\|_{\square, S_i \times S_j} \le d_{ij} + |w_{ij} - w| \le 2\varepsilon/3, \quad 1 \le i < j \le r.$$

On the other hand, trivially, for every i,

$$\|W - w\|_{\square, S_i \times S_i} \le \sup |W - w| \le 1.$$

Let $B := \bigcup_{i=1}^{r} S_i$. Then $\lambda(B) = r/k \ge r/K$ and, recalling that the sets S_i have the same measure,

$$\begin{aligned} \|W - w\|_{\square, B} &\le r^{-2} \sum_{i,j=1}^{r} \|W - w\|_{\square, S_i \times S_j} \le r^{-2}\Big(r(r-1)\frac{2\varepsilon}{3} + r \cdot 1\Big) \\ &< \frac{2\varepsilon}{3} + \frac{1}{r} \le \varepsilon. \end{aligned}$$

(ii): $k < 2M$. We simple take $B = S_1$ and $W = w_{11}$. Then $\lambda(B) = 1/k > 1/(2M)$, and (A.1) implies

$$\|W - w\|_{\square, B} \le \lambda(B)^{-2}\|W - w\|_{\square} \le \lambda(B)^{-2}\frac{\varepsilon}{4M^2} < \varepsilon.$$

This completes the proof of Lemma A.1. □

Proof of Lemma 5.3. We may assume that $\gamma = 0$.

For $\varepsilon > 0$ and $\eta > 0$, let

$$E_{\varepsilon,\eta} := \left\{ (x, y) \in (0, 1)^2 : (2\varepsilon)^{-2} \int_{|x'-x|, |y'-y| < \varepsilon} |W(x', y') - W(x, y)| \, \mathrm{d}x' \, \mathrm{d}y' < \eta \right\}. \tag{A.3}$$

The Lebesgue differentiation theorem says that a.e. $(x, y) \in \bigcap_{\eta} \bigcup_{\varepsilon} E_{\varepsilon,\eta}$; in other words, a.e. $(x, y) \in E_{\varepsilon,\eta}$ for every $\eta > 0$ and all sufficiently small $\varepsilon > 0$ (depending on x, y and η). For $\eta > 0$ and $n \ge 1$, we can thus find $\varepsilon_1 = \varepsilon_1(\eta, n) \in (0, 1/n)$ such that $\lambda\big(E_{\varepsilon_1(\eta,n),\eta}\big) > 1 - 2^{-n}$.

For $n \ge 1$, let $\delta_n := \delta(1/n)$ be as in Lemma A.1 with $\varepsilon = 1/n$, and let $\eta_n := \delta_n^2/n$, $\varepsilon_2(n) := \varepsilon_1(\eta_n, n)$ and $E_n := E_{\varepsilon_2(n),\eta_n}$. Then $\lambda\big(E_n\big) > 1 - 2^{-n}$, so if $\tilde{E} := \bigcup_{n=1}^{\infty} \bigcap_{\ell=n}^{\infty} E_\ell$, then $\lambda(\tilde{E}) = 1$. Let $E := \tilde{E} \cup \{(x, x) : x \in [0, 1]\}$.

For $x \in (0, 1)$ and n so large that $A_n(x) := (x - \varepsilon_2(n), x + \varepsilon_2(n)) \subset (0, 1)$, use Lemma A.1 to find $w_n(x)$ and a set $B_n(x) \subseteq A_n(x)$ with $\lambda(B_n(x)) \ge \delta_n \lambda(A_n(x)) = 2\delta_n \varepsilon_2(n)$ such that

$$\|W - w_n(x)\|_{\square, B_n(x)} \le 1/n. \tag{A.4}$$

If $(x, y) \in E$ and $x \neq y$, then $(x, y) \in \tilde{E}$ so for all large n, $(x, y) \in E_n = E_{\varepsilon_2(n),\eta_n}$, and thus, by (A.3),

$$\begin{aligned}
&\int_{B_n(x)\times B_n(y)} |W(x', y') - W(x, y)| \,\mathrm{d}\lambda_{B_n(x)}(x')\,\mathrm{d}\lambda_{B_n(y)}(y') \\
&\leq (2\delta_n\varepsilon_2(n))^{-2} \int_{A_n(x)\times A_n(y)} |W(x', y') - W(x, y)| \,\mathrm{d}x'\,\mathrm{d}y' \\
&< \delta_n^{-2}\eta_n = 1/n.
\end{aligned} \tag{A.5}$$

Let χ be a Banach limit, i.e., a multiplicative linear functional on ℓ^∞ such that $\chi((a_n)_1^\infty) = \lim_{n\to\infty} a_n$ if the limit exists. Now define

$$W'(x, y) := \begin{cases} \chi\big((w_n(x))_n\big), & y = x, \\ W(x, y), & y \neq x. \end{cases} \tag{A.6}$$

Note that W' is a graphon and a version of W. (Lebesgue measurability is immediate, since the diagonal is a null set.)

Assume for the rest of the proof that $x_1, \dots, x_m \in (0, 1)^m$ with $(x_i, x_j) \in E$ for all i and j. For sufficiently large n, (A.4) holds for all x_i and (A.5) holds for all pairs (x_i, x_j) with $x_i \neq x_j$. Thus, if $x_i = x_j$, by (A.4),

$$\|W - w_n(x_i)\|_{\square,B_n(x_i)\times B_n(x_j)} \leq 1/n, \tag{A.7}$$

and if $x_i \neq x_j$, by (A.5), since the cut norm is at most the L^1 norm,

$$\|W - W(x_i, x_j)\|_{\square,B_n(x_i)\times B_n(x_j)} \leq 1/n. \tag{A.8}$$

For notational convenience, we define the constants

$$w_{ij,n} := \begin{cases} w_n(x_i), & x_i = x_j, \\ W(x_i, x_j), & x_i \neq x_j, \end{cases} \tag{A.9}$$

and let $B_{ni} := B_n(x_i)$. Thus, (A.7) and (A.8) say that for all $i, j \in [m]$,

$$\|W - w_{ij,n}\|_{\square,B_{ni}\times B_{nj}} \leq 1/n. \tag{A.10}$$

We extend the definition of Φ_W in (5.1) to families $(W_{ij})_{1\leq i<j\leq m}$ of functions and write

$$\Phi[(W_{ij})](y_1, \dots, y_m) := \Phi\big((W_{ij}(x_i, x_j))_{i<j}\big).$$

A standard argument shows that, for $|W_{ij}| \leq 1$, say, for all i and j, and any sets $B_1, \dots, B_m \subseteq [0, 1]$ with positive measures, the mapping

$$(W_{ij}) \mapsto \Phi[(W_{ij}); B_1, \dots, B_m] := \int_{B_1\times\cdots\times B_m} \Phi[(W_{ij})](y_1, \dots, y_m)\,\mathrm{d}\lambda_{B_1}(y_1)\dots\mathrm{d}\lambda_{B_m}(y_m)$$

is Lipschitz in cut norm, in each variable separately; by linearity it suffices to consider the case when Φ is a monomial (and thus $\Phi_W = \Psi_{F,W}$ for some graph F), and this result is then explicit in [2, Proof of Lemma 2.2]; see also [4]. Thus, by (A.10), recalling that each $w_{ij,n}$ here is a constant,

$$\Phi[(W); B_{n1}, \dots, B_{nm}] - \Phi((w_{ij,n})_{i<j}) = O(1/n). \tag{A.11}$$

On the other hand, $\Phi[(W)](y_1, \dots, y_m) = \Phi_W(y_1, \dots, y_m) = \gamma = 0$ a.e., by assumption, and thus $\Phi[(W); B_{n1}, \dots, B_{nm}] = 0$. Consequently, (A.11) yields

$$\Phi((w_{ij,n})_{i<j}) = O(1/n). \tag{A.12}$$

Apply the Banach limit χ to (A.12). With $z_{ij} := \chi((w_{ij,n})_{i<j})$ we obtain, recalling that Φ is a polynomial,

$$\Phi((z_{ij})_{i<j}) = 0. \tag{A.13}$$

If $x_i \neq x_j$, then, by (A.9), $w_{ij,n} = W(x_i, x_j)$ for all n, and thus $z_{ij} = W(x_i, x_j) = W'(x_i, x_j)$; see (A.6). If $x_i = x_j$, then (A.9) shows that $w_{ij,n} = w_n(x_i)$, and thus, using (A.6), $z_{ij} = \chi((w_n(x_i))_n) = W'(x_i, x_j)$. Consequently, $z_{ij} = W'(x_i, x_j)$ for all (i, j), and (A.13) can be written $\Phi_{W'}(x_1, \dots, x_m) = 0$, as asserted. (In order to avoid any worry of edge effects, we have considered $x_i \in (0, 1)$ only. For completeness, we, trivially, may define $W'(0, 0) := W'(1, 1) := W'(\frac{1}{2}, \frac{1}{2})$.) □

Finally, we mention another technical problem, which might be of interest in some applications:

Problem A.4. The version W' in Lemma 5.3 is Lebesgue measurable. Can W' always be chosen to be Borel measurable?

(The construction in the proof above, using a Banach limit, does not seem to guarantee Borel measurability.)

References

[1] B. Bollobás, S. Janson, O. Riordan, The cut metric, random graphs, and branching processes, Random Structures Applications (in press).
[2] B. Bollobás, O. Riordan, Metrics for sparse graphs, in: Surveys in Combinatorics 2009, in: LMS Lecture Notes Series, vol. 365, Cambidge Univ. Press, 2009, pp. 211–287.
[3] C. Borgs, J. Chayes, L. Lovász, Moments of two-variable functions and the uniqueness of graph limits, Geom. Funct. Anal. 19 (6) (2010) 1597–1619.
[4] C. Borgs, J.T. Chayes, L. Lovász, V.T. Sós, K. Vesztergombi, Convergent sequences of dense graphs I: subgraph frequencies, metric properties and testing, Adv. Math. 219 (2008) 1801–1851.
[5] C. Borgs, J.T. Chayes, L. Lovász, V.T. Sós, K. Vesztergombi, Convergent sequences of dense graphs II: multiway cuts and statistical physics. Preprint, 2007. http://research.microsoft.com/~borgs/.
[6] F.R.K. Chung, R.L. Graham, Maximum cuts and quasirandom graphs, in: Random Graphs, Vol. 2 (Poznań, 1989), Wiley, New York, 1992, pp. 23–33.
[7] F.R.K. Chung, R.L. Graham, R.M. Wilson, Quasi-random graphs, Combinatorica 9 (4) (1989) 345–362.
[8] P. Diaconis, S. Janson, Graph limits and exchangeable random graphs, Rendiconti di Matematica 28 (2008) 33–61.
[9] P. Diaconis, S. Holmes, S. Janson, Threshold graph limits and random threshold graphs, Internet Math. 5 (2009) 267–318.
[10] G. Elek, B. Szegedy, A measure-theoretic approach to the theory of dense hypergraphs. arXiv:0810.4062v2.
[11] D.H. Gottlieb, A certain class of incidence matrices, Proc. Amer. Math. Soc. 17 (1966) 1233–1237.
[12] R.L. Graham, B.L. Rothschild, J.H. Spencer, Ramsey Theory, Wiley, New York, 1980.
[13] L. Lovász, V.T. Sós, Generalized quasirandom graphs, J. Combin. Theory B. 98 (1) (2008) 146–163.
[14] L. Lovász, B. Szegedy, Limits of dense graph sequences, J. Combin. Theory B 96 (2006) 933–957.
[15] L. Lovász, B. Szegedy, Szemerédi's lemma for the analyst, Geom. Funct. Anal. 17 (1) (2007) 252–270.
[16] A. Shapira, Quasi-randomness and the distribution of copies of a fixed graph, Combinatorica 28 (2008) 735–745.
[17] A. Shapira, R. Yuster, The effect of induced subgraphs on quasi-randomness, in: Proc. 19th Annual ACM–SIAM Symposium on Discrete Algorithms, SODA, ACM Press, 2008, pp. 789–798 (Preliminary version).
[18] M. Simonovits, V.T. Sós, Szemerédi's partition and quasirandomness, Random Structures Algorithms 2 (1) (1991) 1–10.
[19] M. Simonovits, V.T. Sós, Hereditarily extended properties, quasi-random graphs and not necessarily induced subgraphs, Combinatorica 17 (4) (1997) 577–596.
[20] M. Simonovits, V.T. Sós, Hereditary extended properties, quasi-random graphs and induced subgraphs, Combin. Probab. Comput. 12 (3) (2003) 319–344.
[21] E.M. Stein, Singular Integrals and Differentiability Properties of Functions, Princeton University Press, Princeton, NJ, 1970.
[22] A. Thomason, Pseudorandom graphs, in: Random Graphs'85 (Poznań, 1985), North-Holland, Amsterdam, 1987, pp. 307–331.
[23] A. Thomason, Random graphs, strongly regular graphs and pseudorandom graphs, in: Surveys in Combinatorics 1987 (New Cross, 1987), in: London Math. Soc. Lecture Note Ser., vol. 123, Cambridge Univ. Press, Cambridge, 1987, pp. 173–195.
[24] R. Yuster, Quasi-randomness is determined by the distribution of copies of a fixed graph in equicardinal large sets, in: Proc. 12th International Workshop on Randomization and Computation, RANDOM, Springer, 2008, pp. 596–601.